

CCGACACGCTGGGTGCGAGGCGAAGGTGACCGGGGACCGAGCATTTGAGATCTGCTCGGTAGA
CCTGGTGCAACCAACCACCATGTTGGCTGCAAGGCTGGTGTGTCTCCGGACACTACCTTCTAGG
GTTTTCCACCCAGCTTTACCAAGGCCTCCCCTGTTGTGAAGAATTCCATCACGAAGAATCA
ATGGCTGTTAACACCTAGCAGGGAATATGCCACCAAAACAAGAATTGGGATCCGGCGTGGGA
GAACTGGCCAAGAAGCTCAAAGAGGCAGCATTGGAACCATCGATGGAAAAAATATTTAAAATT
GATCAGATGGGAAGATGGTTTGTGTGCTGGAGGGCTGCTGTTGCTCTTGAGCATTGTGCTA
CTATGGCTTGGGACTGTCTAATGAGATTGGAGCTATTGAAAAGGCTGTAATTTGGCCTCAGT
ATGTCAAGGATAGAATTCAATCCACCTATATGTACTTAGCAGGGAGTATTGGTTTAACAGCT
TTGTCTGCCATAGCAATCAGCAGAACGCCTGTTCTCATGAACCTCATGATGAGAGGCTCTTG
GGTGACAATTGGTGTGACCTTTGCAAGCATGGTTGGAGCTGGAATGCTGGTACGATCAATAC
CATATGACCAGAGCCCCAGGCCAAAGCATCTTGCTTGGTTGCTACATTCTGGTGTGATGGGT
GCAGTGGTGGCTCCTCTGACAATATTAGGGGGTCTCTTCTCATCAGAGCTGCATGGTACAC
AGCTGGCATTGTGGGAGGCCTCTCCACTGTGGCCATGTGTGCGCCAGTGAAAAGTTTCTGA
ACATGGGTGCACCCCTGGGAGTGGGCCTGGGTCTCGTCTTTGTGTCTCATTGGGATCTATG
TTTCTTCCACCTACCACCGTGGCTGGTGCCACTCTTTACTCAGTGGCAATGTACGGTGGATT
AGTTCTTTTTCAGCATGTTCCCTTCTGTATGATACCAGAAAGTAATCAAGCGTGCAGAAGTAT
CACCAATGTATGGAGTTCAAAAATATGATCCCATTAACCTCGATGCTGAGTATCTACATGGAT
ACATTAAATATATTTATGCGAGTTGCAACTATGCTGGCAACTGGAGGCAACAGAAAGAAATG
AAGTGACTCAGCTTCTGGCTTCTCTGCTACATCAAATATCTTGTTAATGGGGCAGATATGC
ATTAAATAGTTTTGTACAAGCAGCTTTCGTTGAAGTTTAGAAGATAAGAAACATGTCATCATA
TTTAAATGTTCCGGTAATGTGATGCCTCAGGTCTGCCTTTTTTCTGGAGAATAAATGCAGT
AATCCTCTCCCAAATAAGCACACACATTTCAATTCTCATGTTTGAAGTATTTTAAATGTT
TTGGTGAATGTGAAAACATAAGTTTTGTGTCATGAGAATGTAAGTCTTTTTTCTACTTTAAAA
TTTAGTAGGTTTCACTGAGTAACTAAAATTTAGCAAACCTGTGTTTGCATATTTTTTTGGAGT
TCAGAATATTGTAATTAATGTCATAAGTGATTTGGAGCTTTGGTAAAGGACCAGAGAGAAG
GAGTCACCTGCAGTCTTTTTGTTTTTTTAAATACTTAGAACTTAGCACTTGTGTTATTGATTA
GTGAGGAGCCAGTAAGAAACATCTGGGTATTTGGAAACAAGTGGTCATTGTTACATTTCATTT
GCTGAACTTAACAAAACCTGTTTCATCCTGAAACAGGCACAGGTGATGCATTCTCCTGCTGTTG
CTTCTCAGTGCTCTCTTTCCAATATAGATGTGGTCATGTTTGACTTGTACAGAATGTTAATC
ATACAGAGAATCCTTGATGGAATTATATATGTGTGTTTTACTTTTTGAATGTTACAAAAGGAA
ATAACTTTAAAACTATTCTCAAGAGAAAATATTCAAAGCATGAAATATGTTGCTTTTTTCCAG
AATACAAACAGTATACTCATG

FIGURE 2

MLAARLVCLRTLPSRVFHPAFTKASPVVKNSITKNQWLLTPSREYATKTRIGIRRGRTGQEL
KEAALEPSMEKIFKIDQMGRWFVAGGAAVGLGALCYGLGLSNEIGAIEKAVIWPQYVKDRI
HSTYMYLAGSIGLTALSAIAISRTPVLMNFMMRGSWVTIGVTFAAMVGAGMLVRSIPYDQSP
GPKHLAWLLHSGVMGAVVAPLTLGGPLLIRAAWYTAGIVGGLSTVAMCAPSEKFLNMGAPL
GVGLGLVVFVSSLGSMFLPPTTVAGATLYSVAMYGGLVLFSMFLLYDTQKVIKRAEVSPMYGV
QKYDPINSMLSIYMDTLNIFMRVATMLATGGNRKK

09989727 441904
T06111 22768880

FIGURE 3

GAAGGCTGCCTCGCTGGTCCGAATTCGGTGGCGCCACGTCCGCCCCGTCTCCGCCTTCTGCAT
 CGCGGCTTCGGCGGCTTCCACCTAGACACCTAACAGTCGCGGAGCCGGCCGCGTCGTGAGGG
 GGTCGGCACGGGGAGTCGGGCGGTCTTGTGCATCTTGGCTACCTGTGGGTGCAAG**ATG**TCGG
 ACATCGGAGACTGGTTCAGGAGCATCCCGGCGATCACGCGCTATTGGTTCGCGGCCACCGTC
 GCCGTGCCCTTGGTCGGCAAACCTCGGCCTCATCAGCCCCGGCCTACCTCTTCCCTCTGGCCCCGA
 AGCCTTCCTTTATCGCTTTCAGATTTGGAGGCCAATCACTGCCACCTTTTATTTCCCTGTGG
 GTCCAGGAACCTGGATTTCTTTATTTGGTCAATTTATATTTCTTATATCAGTATTCTACGCGA
 CTTGAAACAGGAGCTTTTGATGGGAGGCCAGCAGACTATTTATTCATGCTCCTCTTTAACTG
 GATTTGCATCGTGATTACTGGCTTAGCAATGGATATGCAGTTGCTGATGATTCTCTGATCA
 TGTCAGTACTTTATGTCTGGGCCAGCTGAACAGAGACATGATTGTATCATTTTTGGTTTGA
 ACACGATTTAAGGCCTGCTATTTACCCTGGGTATCCTTGGATTCAACTATATCATCGGAGG
 CTCGGTAATCAATGAGCTTATTGGAATCTGGTTGGACATCTTTATTTTTTCTAATGTTCA
 GATACCCAATGGACTTGGGAGGAAGAAATTTTCTATCCACACCTCAGTTTTTGTACCGCTGG
 CTGCCCAGTAGGAGAGGAGGAGTATCAGGATTTGGTGTGCCCCCTGCTAGCATGAGGCGAGC
 TGCTGATCAGAATGGCGGAGGCGGGAGACACAACCTGGGGCCAGGGCTTTTCTGACTTGAGACC
 AG**TGA**AGGGGGCGGCCTCGGGCAGCCGCTCCTCTCAAGCCACATTTCTCCAGTGCTGGGTG
 CACTTAACAACCTGCGTTCTGGCTAACACTGTTGGACCTGACCCACACTGAATGTAGTCTTTC
 AGTACGAGACAAAGTTTCTTAAATCCCGAAGAAAAATATAAGTGTTCCACAAGTTTACGAT
 TCTCATTCAAGTCCTTACTGCTGTGAAGAACAAATACCAACTGTGCAAAATGCAAACTGAC
 TACATTTTTTTGGTGTCTTCTCTTCTCCCCCTTTCCGTCTGAATAATGGGTTTTAGCGGGTCT
 AATCTGCTGGCATTGAGCTGGGGCTGGGTCAACAAACCTTCCCAAAGGACCTTATCTCTT
 TCTTGACACATGCCTCTCTCCCACTTTTCCCAACCCCCACATTTGCAACTAGAAAAAGTTG
 CCCATAAAATTGCTCTGCCCTTGACAGTTCTGTTATTTATTGACTTTTGCCAAGGCTGGT
 ACAACAATCATATTCACGTTATTTTCCCCCTTTTGGTGGCAGAACTGTTACCAATAGGGGGAG
 AAGACAGCCACGGATGAAGCGTTTCTCAGCTTTTGGAAATTGCTTCGACTGACATCCGTTGTT
 AACCGTTTGCCACTCTTCAGATATTTTTTATAAAAAAGTACCACTGAGTTCATGAGGGCCA
 CAGATTGGTTATTAATGAGATACGAGGGTTGGTGTGGGTGTTTGTTCCTGAGCTAAGTGA
 TCAAGACTGTAGTGGAGTTGCAGCTAACATGGGTTAGGTTTAAACCATGGGGGATGCACCC
 TTTGCGTTTCATATGTAGCCCTACTGGCTTTGTGTAGCTGGAGTAGTTGGGTGCTTTGTGT
 TAGGAGGATCCAGATCATGTTGGCTACAGGGAGATGCTCTCTTTGAGAGGTCCTGGGCATTG
 ATTCCCATTTCATCTCATTTCTGGATATGTGTTTATTGAGTAAAGGAGGAGAGACCCTCATA
 CGCTATTTAAATGTCACCTTTTTTGCCCTATCCCCCGTTTTTTGGTTCATGTTTCAATTAATTGT
 GAGGAAGGCGCAGCTCCTCTCTGCACGTAGATCATTTTTTAAAGCTAATGTAAGCACATCTA
 AGGGAATAACATGATTTAAGGTTGAAATGGCTTTAGAATCATTTGGGTTTGAGGGTGTGTTA
 TTTTGAGTCATGAATGTACAAGCTCTGTGAATCAGACCAGCTTAAATACCCACACCTTTTTT
 TCGTAGGTGGGCTTTTCTATCAGAGCTTGGCTCATAACCAAATAAAGTTTTTTGAAGGCCA
 TGGCTTTTTCACACAGTTATTTTATTTTATGACGTTATCTGAAAGCAGACTGTTAGGAGCAGT
 ATTGAGTGGCTGTCACACTTTGAGGCAACTAAAAAGGCTTCAAACGTTTTTGATCAGTTTCTT
 TTCAGGAAACATTGTGCTCTAACAGTATGACTATTCTTTCCCCCACTCTTAAACAGTGTGAT
 GTGTGTTATCCTAGGAAATGAGAGTTGGCAAACAACCTTCTCATTTTGAATAGAGTTTGTGTG
 TACTTCTCCATATTTAATTTATATGATAAAATAGGTGGGGAGAGTCTGAACCTTAACTGTCA
 TGTTTTGTTGTTTATCTGTGGCCACAATAAAGTTTACTTGTAAAATTTTAGAGGCCATTACT
 CCAATTATGTTGCACGTACACTCATTGTACAGGCGTGGAGACTCATTGTATGTATAAGAATA
 TTTCTGACAGTGAGTGACCCGGAGTCTCTGGTGTACCCTCTTACCAGTCAGCTGCCTGCGAG
 CAGTCATTTTTTCTTAAAGGTTTACAAGTATTTAGAACTTTTCAGTTCAGGGCAAAATGTTT
 ATGAAGTTATTCCTCTTAAACATGGTTAGGAAGCTGATGACGTTATTGATTTTGTCTGGATT
 ATGTTTTCTGGAATAATTTTACCAAAACAAGCTATTTGAGTTTTGACTTGACAAGGCAAAACA
 TGACAGTGGATTCTCTTTACAAATGAAAAAATAATCCTTATTTTGTATAAAGGACTTCCC
 TTTTTGTAACTAATCCTTTTTTATTGGTAAAAATTGTAAATTAAATGTGCAACTTG

FIGURE 4

MSDIGDWFRSIPAITRYWFAATVAVPLVGKLGLISPAYLFLWPEAFLYRFQIWRPITATFYF
PVGPGTGFLYLVNLYFLYQYSTRLETGAFDGRPADYLFMLLFNWICIVITGLAMDMQLLMIP
LIMSVLYVWAQLNRDMIVSFWFGTRFKACYLPWVILGFNYIIIGGSVINELIGNLVGHLYFFL
MFRYPMDLGGRNFLSTPQFLYRWLPSRRGGVSGFGVPPASMRRAADQNGGGGRHNWGQGFRL
GDQ

Transmembrane domain:

amino acids 98-116, 152-172

N-myristoylation site.

amino acids 89-95, 168-174, 176-182, 215-221, 221-227, 237-243

Glycosaminoglycan attachment site.

amino acids 218-222

FIGURE 5

GGGGCCGCGGTCTAGGGCGGCTACGTGTGTTGCCATAGCGACCATTTTGCATTAACTGGTTG
GTAGCTTCTATCCTGGGGGCTGAGCGACTGCGGGGCCAGCTCTTCCCCTACTCCCTCTCGGCT
CCTTGTGGCCCAAAGGCCTAACCGGGGTCCGGCGGTCTGGCCTAGGGATCTTCCCCGTTGCC
CCTTTGGGGCGGG**ATG**GCTGCGGAAGAAGAAGACGAGGTGGAGTGGGTAGTGGAGAGCATCG
CGGGGTTCCTGCGAGGCCCAGACTGGTCCATCCCCATCTTGGACTTTGTGGAACAGAAATGT
GAAGTTAACTGCAAAGGAGGGCATGTGATAACTCCAGGAAGCCCAGAGCCGGTGATTTTGGT
GGCCTGTGTTCCCCTTGTTTTTGGATGATGAAGAAGAAAGCAAATTGACCTATACAGAGATTC
ATCAGGAATACAAAGAACTAGTTGAAAAGCTGTTAGAAGGTTACCTCAAAGAAATTGGAATT
AATGAAGATCAATTTCAAGAAGCATGCACTTCTCCTCTTGCAAAGACCCATACATCACAGGC
CATTTTGCAACCTGTGTTGGCAGCAGAAGATTTTACTATCTTTAAAGCAATGATGGTCCAGA
AAAACATTGAAATGCAGCTGCAAGCCATTTCGAATAATTCAAGAGAGAAATGGTGTATTACCT
GACTGCTTAACCGATGGCTCTGATGTGGTCAGTGACCTTGAACACGAAGAGATGAAAATCCT
GAGGGAAGTTCTTAGAAAATCAAAGAGGAATATGACCAGGAAGAAGAAAGGAAGAGGAAAA
AACAGTTATCAGAGGCTAAACAGAAGAGCCACAGTGCATTCCAGTGAAGCTGCAATAATG
AATAATTCCTAAGGGGATGGTGAACATTTTGCACACCCACCCTCAGAAGTTAAAATGCATTT
TGCTAATCAGTCAATAGAACCCTTTGGGAAGAAAAGTGGAAAGGTCTGAAACTTCCTCCCTCC
CACAAAAAGGCCTGAAGATTCTTGGCTTAGAGCATGCGAGCATTGAAGGACCAATAGCAAAC
TTATCAGTACTTGGAACAGAAGAAGCTTCGGCAACGAGAACACTATCTCAAGCAGAAGAGAGA
TAAGTTGATGTCCATGAGAAAGGATATGAGGACTAAACAGATACAAAATATGGAGCAGAAAG
GAAAACCCACTGGGGAGGTAGAGGAAATGACAGAGAAACCAGAAATGACAGCAGAGGAGAAG
CAAACATTACTAAAGAGGAGATTGCTTGCAGAGAACTCAAAGAAGAAGTTATTAATAAG**TA**
ATAATTAAGAACAATTTAACAAAATGGAAGTTCAAATTGTCTTAAAAATAAATTATTTAGTC
CTTACACTG

FIGURE 6

MAAEEEEDEVEWVVESIAGFLRGPDWSIPILDFVEQKCEVNCKGGHVITPGSPEPVILVACVP
LVFDDEEESKLTYTEIHQEYKELVEKLLLEGYLKEIGINEDQFQEAactsPLAKTHTSQAILQP
VLAAEDFTIFKAMMVQKNIEMQLQAIRIIQERNGLVPDCLTDGSDVVSdleHEEMKILREVL
RKSKEEYDQEEERKRKKQLSEAKTEEPTVHSSEAAIMNNSQGDGEHFAHPPSEVKMHFANQS
IEPLGRKVERSETSSLPQKGLKIPGLEHASIEGPIANLSVLGTEELRQREHYLKQKRDKLMS
MRKDMRTKQIQNMEQKGKPTGEVEEMTEKPEMTAEKQTLLKRRLLAEKLKEEVINK

N-glycosylation sites.

amino acids 224-228, 246-250, 285-289

N-myristoylation site.

amino acids 273-279

Amidation site.

amino acids 252-256

Cytosolic fatty-acid binding proteins.

amino acids 78-108

FIGURE 7

GGGCACAGCACATGTGAAGTTTTTGATGATGAAGAAGAAAGCAAATTGACCTATACAGAGAT
TCATCAGGAATACAAAGAACTAGTTGAAAAGCTGTTAGAAGGTTACCTCAAAGAAATTGGAA
TTAATGAAGATCAATTTCAAGAAGCATGCACTTCTCCTCTTGCAAAGACCCATACATCACAG
GCCATTTTTTGCAACCTGTGTTGGCAGCAGAAGATTTTACTATCTTTAAAGCAATGATGGTCC
AGAAAAACATTGAAATGCAGCTGCAAGCCATTCGAATAATTCAAGAGAGAAATGGTGTATTA
CCTGACTGCTTAACCGATGGCTCTGATGTGGTCAGTGACCTTGAACACGAAGAGATGAAAAT
CCTGAGGGAAGTTCTTAGAAAATCAAAAGAGGAATATGACCAGGAA

FIGURE 8

GCGTGGTTTTTGTCTGCAATAGGCGGCTTAGAGGGAGGGGCTTTTTTCGCTATACCTACTG
TAGCTTCTCCACGTATGGACCCTAAAGGCTACTGCTGCTACTACGGGGCTAGACAGTTACTG
TCTCAGCTCTAGGATGTGCGTTCTTCCACTAGAAGCTCTTCTGAGGGAGGTAATTAAAAAAC
AGTGGAATGGAAAAACAGTGCTGTAGTCATCCTGTAATATGCTCCTTGTCAACAATGTATAC
ATTCCCTGCTAGGTGCCATATTCATTGCTTTAAGCTCAAAGTCGCATCTTACTAGTGAAGTATT
CTGCCAATGAAGAAAAACAAGTATGATTATCTTCCAACCTACTGTGAATGTGTGCTCAGAACTG
GTGAAGCTAGTTTTCTGTGTGCTTGTGTCAATTCTGTGTTATAAAGAAAGATCATCAAAGTAG
AAATTTGAAATATGCTTCCTGGAAGGAATTCTCTGATTTTCATGAAGTGGTCCATTTCCTGCCT
TTCTTTATTTTCCTGGATAACTTGATTGTCTTCTATGTCCTGTCCTATCTTCAACCAGCCATG
GCTGTTATCTTCTCAAATTTTAGCATTATAACAACAGCTCTTCTATTTCAGGATAGTGCTGAA
GAGGCGTCTAAACTGGATCCAGTGGGCTTCCCTCCTGACTTTATTTTTGTCTATTGTGGCCT
TGACTGCCGGGACTAAAACCTTACAGCACAACTTGGCAGGACGTGGATTTTCATCACGATGCC
TTTTTCAGCCCTTCCAATTCCTGCCTTCTTTTCAGAAGTGAGTGTCCCAGAAAAGACAATTG
TACAGCAAAGGAATGGACTTTTCCCTGAAGCTAAATGGAACACCACAGCCAGAGTTTTTCAGTC
ACATCCGTCTTGGCATGGGCCATGTTCTTATTATAGTCCAGTGTTTTATTTCTTCAATGGCT
AATATCTATAATGAAAAGATACTGAAGGAGGGGAACCAGCTCACTGAAAGCATCTTCATACA
GAACAGCAAACCTCTATTTCTTTGGCATTCTGTTAATGGGCTGACTCTGGGCCTTCAGAGGA
GTAACCGTGATCAGATTAAGAACTGTGGATTTTTTTTATGGCCACAGTGCATTTTCAGTAGCC
CTTATTTTTGTAACTGCATTCCAGGGCCTTTCAGTGGCTTTCATTCTGAAGTTCCTGGATAA
CATGTTCCATGTCTTGATGGCCCAGGTTACCACTGTCAATTATCACAAACAGTGTCTGTCTCTGG
TCTTTGACTTCAGGCCCTCCCTGGAATTTTTCTTGGAAAGCCCCATCAGTCCTTCTCTCTATA
TTTATTTATAATGCCAGCAAGCCTCAAGTTCGGAATACGCACCTAGGCAAGAAAGGATCCG
AGATCTAAGTGGCAATCTTTGGGAGCGTTCAGTGGGGATGGAGAAGAACTAGAAAGACTTA
CCAAACCCAAGAGTGATGAGTCAGATGAAGATACTTTCTAACTGGTACCCACATAGTTTGCA
GCTCTCTTGAACCTTATTTTCACATTTTCAGTGTTTGTAATATTTATCTTTTCACTTTGATA
AACCAGAAATGTTTCTAAATCCTAATATTTCTTGCATATATCTAGCTACTCCCTAAATGGTT
CCATCCAAGGCTTAGAGTACCCAAAGGCTAAGAAATTTCTAAAGAACTGATACAGGAGTAACA
ATATGAAGAATTCATTAATATCTCAGTACTTGATAAATCAGAAAGTTATATGTGCAGATTAT
TTTCCTTGGCCTTCAAGCTTCCAAAAAACTTGTAATAATCATGTTAGCTATAGCTTGTATAT
ACACATAGAGATCAATTTGCCAAATATTCACAATCATGTAGTTCTAGTTTACATGCCAAAGT
CTTCCCTTTTTAACATTATAAAAGCTAGGTTGTCTCTTGAATTTTGAGGCCCTAGAGATAGT
CATTTTGCAAGTAAAGAGCAACGGGACCCTTTCTAAAAACGTTGGTTGAAGGACCTAAATAC
CTGGCCATACCATAGATTTTGGGATGATGTAGTCTGTGCTAAATATTTTGCTGAAGAAGCAGT
TTCTCAGACACAACATCTCAGAATTTTAATTTTTAGAAATTCATGGGAAATTGGATTTTTGT
AATAATCTTTTGATGTTTTAAACATTGGTTCCCTAGTCACCATAGTTACCACTTGTATTTTA
AGTCATTTAAACAAGCCACGGTGGGGCTTTTTTCTCCTCAGTTTGAGGAGAAAAATCTTGAT
GTCATTACTCCTGAATTATTACATTTTGGAGAATAAGAGGGCATTATTTTATTAGTTACT
AATTCAAGCTGTGACTATTGTATATCTTTCCAAGAGTTGAAATGCTGGCTTCAGAATCATAC
CAGATTGTCAAGTGAAGCTGATGCCTAGGAACTTTTAAAGGGATCCTTTCAAAGGATCACTT
AGCAAACACATGTTGACTTTTAACTGATGTATGAATATTAATACTCTAAAAATAGAAAGACC
AGTAATATATAAGTCACTTTACAGTGCTACTTCACACTTAAAGTGATGGTATTTTTTCATG
GTATTTTGCATGCAGCCAGTTAACTCTCGTAGATAGAGAAGTCAGGTGATAGATGATATTAA
AAATTAGCAAACAAAAGTGACTTGCTCAGGGTCATGCAGCTGGGTGATGATAGAAGAGTGGG
CTTTAACTGGCAGGCCTGTATGTTTACAGACTACCATACTGTAAATATGAGCTTTATGGTGT
CATTTCTCAGAACTTATACATTTCTGCTCTCCTTTCTCCTAAGTTTCATGCAGATGAATATA
AGGTAATATACTATTATATAATTCATTTGTGATATCCACAATAATATGACTGGCAAGAATTG
GTGGAAATTTGTAATTAAATAATTATTAAACCT

FIGURE 9

MEKQCCSHPVICSLSTMYTFLLGAIFIALSSSRILLVKYSANEENKYDYLPTTVNVCSELVK
LVFCVLVSFCVIKKDHQSRNLKYASWKEFSDFMKWSIPAFLYFLDNLIVFYVLSYLQPAMAV
IFSNFSIITTALLFRIVLKRRLNWIQWASLLTLFLSIVALTAGTKTLQHNLAGRGFHHDAFF
SPSNSCLLFRSECPRKDNCTAKEWTFPEAKWNTTARVFSHIRLGMGHVLIIVQCFISSMANI
YNEKILKEGNQLTESIFIQNSKLYFFGILFNGLTLGLQRSNRDQIKNCGFFYGHSAFSVALI
FVTAFQGLSVAFILKFLDNMFHVLMAQVTTVIITTVSVLVFDFRPSLEFFLEAPSVLLSIFI
YNASKPQVPEYAPRQERIRDLSGNLWERSSSGDGEELERLTKPKSDESDETF

Transmembrane domains:

amino acids 16-36 (type II), 50-74, 147-168, 229-250, 271-293,
298-318, 328-368

N-glycosylation sites.

amino acids 128-132, 204-208, 218-222, 374-378

Glycosaminoglycan attachment site.

amino acids 402-406

N-myristoylation sites.

amino acids 257-263, 275-281, 280-286, 284-290, 317-323

FIGURE 10

CGTGCCTGCGCAATGGGTGTCGGGTCCGCTTTTTCCCAATCCGGACGTAATCGTGGTTTTTG
TTCTGCAATAGGCGGCTTAGAGGGAGGGGCTTTTTCGCCTATACCTACTGTAGCTTCTCCAC
GTATGGACCCTAAAGGCTACTGCTGCTACTACGGGGCTAGACAGTTACTGTCTCAGCTCTAG
GATGTGCGTTCTTCCACTAGAAGCTCTTCTGAGGGAGGTAATTAAAAACAGTGGAATGGAA
AAACAGTGCTGTAGTCATCCTGTAATATGCTCCTTGTCAACAATGTATACATTCCTGCTAGG
TGCCATATTCATTGCTTTAAGCTCAAGTCGCATCTTACTAGTGAAGTATTCTGCCAATGAAG
AAAACAAGTATGATTATCTTCCAACACTACTGTGAATGTGTGCTCAGAACTGGTGAAGCTAGTT
TTCTGTGTGCTTGTGTCATTCTGTGTTATAAAGAAAGATCATCAAAGTAGAAATTTGAAATA
TGCTTCCTGGAAGGAATTCTCTGATTTTCATGAAGTGGTCCATTCCTGCCTTTCTTTATTTCC
TGGATAACTTGATTGTCTTCTATGTCCTGTCCTATCTTCAACCAGCCATGGCTGTTATCTTC
TCAAATTTTAGCATTATAACAACAGCTCTTCTATTTCAGGATAGTGCTGAAGAGGCGTCTAAA
CTGGATCCAGTGGGCTTCCCTCCTGACTTTATTTTTGTCTATTGTGGCCTTGACTGCCGGGA
CTAAAACTTTA

CGGACGCGTGGGCGGACGCGCTGGGCGGACGCGTGGGGCCGGCTTGGTATAGCGCGCGGCGGCC
GTGGCTAAGGCTGCTACGAAGCGAGCTTGGGAGGAGCAGCGGCCCTGCGGGGCAGAGGAGCAT
CCCGTCTACCAGGTCCCAAGCGGCGTGGCCCGCGGGTCATGGCCAAAGGAGAAGGCGCCGAG
AGCGGCTCCGCGGCGGGGCTGCTACCCACCAGCATCCTCCAAAGCACTGAACGCCCGGCCCA
GGTGAAGAAAGAACC GAAAAAGAAGAAACAACAGTTGTCTGTTTGCAACAAGCTTTGCTATG
CACTTGGGGGAGCCCCCTACCAGGTGACGGGCTGTGCCCTGGGTTTCTTCCTTCAGATCTAC
CTATTGG**ATG**TGGCTCAGGTGGGCCCTTTCTCTGCCTCCATCATCCTGTTTGTGGGCCGAGC
CTGGGATGCCATCACAGACCCCCCTGGTGGGCCTCTGCATCAGCAAATCCCCCTGGACCTGCC
TGGGTGCGCTTATGCCCTGGATCATCTTCTCCACGCCCCCTGGCCGTCAATTGCCTACTTCCTC
ATCTGGTTTCGTGCCCCACTTCCCACACGGCCAGACCTATTGGTACCTGCTTTTCTATTGCCT
CTTTGAAACAATGGTCACGTGTTTCCATGTTCCCTACTCGGCTCTCACCATGTTTCATCAGCA
ACCGAGCAGACTGAGCGGGATTCTGCCACCGCCTATCGGATGACTGTGGAAGTGCTGGGCAC
AGTGCTGGGCACGGCGATCCAGGGACAAATCGTGGGCCAAGCAGACACGCCTTGTTCAGG
ACTTCAATAGCTCTACAGTAGCTTACAAAGTGCCAACCATAACATGGCACCACCTTCACAC
AGGGAACGCAAAAGGCATACCTGCTGGCAGCGGGGTCATTGTCTGTATCTATATAATCTG
TGCTGTCACTCCTGATCCTGGGCGTGCGGGAGCAGAGAGAACCCCTATGAAGCCAGCAGTCTG
AGCCAATCGCCTACTTCCGGGGCTACGGCTGGTCA TGAGCCACGGCCCATACATCAAACCTT
ATTACTGGCTTCTCTTACCTCCTTGGCTTTTATGCTGGTGGAGGGGAACCTTTGTCTTGT
TTGCACCTACACCTTGGGCTTCCGCAATGAATTCAGAATCTACTCCTGGCCATCATGCTCT
CGGCCACTTTAACCATTCCCATCTGGCAGTGGTTCTTGACCCGTTTGGCAAGAAGACAGCT
GTATATGTTGGGATCTCATCAGCAGTGCCATTTCTCATCTTGGTGGCCCTCATGGAGAGTAA
CCTCATCATTACATATGCGGTAGCTGTGGCAGCTGGCATCAGTGTGGCAGCTGCCTTCTTAC
TACCTGGTCCATGCTGCCTGATGTATTGACGACTTCCATCTGAAGCAGCCCCACTTCCAT
GGAACCGAGCCCATCTTCTTCTCCTTCTATGTCTTCTTCACCAAGTTTGCCTCTGGAGTGT
ACTGGGCATTTCTACCCTCAGTCTGGACTTTGCAGGGTACCAGACCCGTGGCTGCTCGCAGC
CGGAACGTGTCAAGTTTACACTGAACATGCTCGTGACCATGGCTCCCATAGTTTCTCATCCTG
CTGGGCCTGCTGCTCTTCAAAATGTACCCCATTGATGAGGAGAGGCGGCGGCAGAATAAGAA
GGCCCTGCAGGCACTGAGGGACGAGGCCAGCAGCTCTGGCTGCTCAGAAACAGACTCCACAG
AGCTGGCTAGCATCCTC**TAG**GGCCCCGCCACGTTGCCCGAAGCCACCATGCAGAAGGCCACAG
AAGGGATCAGGACCTGTCTGCCGGCTTGCTGAGCAGCTGGACTGCAGGTGCTAGGAAGGGAA
CTGAAGACTCAAGGAGGTGGCCCAGGACACTTGCTGTGCTCACTGTGGGGCCGGCTGCTCTG
TGGCCTCCTGCCTCCCCCTCTGCCTGCCTGTGGGGCCAAGCCCTGGGGCTGCCACTGTGAATA
TGCCAAGGACTGATCGGGCCTAGCCCGGAACACTAATGTAGAAACCTTTTTTTTACAGAGCC
TAATTAATAACTTAATGACTGTGTACATAGCAATGTGTGTGTATGTATATGTCTGTGAGCTA
TTAATGTTATTAATTTTTCATAAAAGCTGGAAAGC

12/330

FIGURE 12

MWLRWALS LPPSSCLWAEPGMPSQTPWWASASANPPGPAWVALCPGSSSPRPWPSLPTSSSG
SCPTSHTARPIGT CFSIASLKQWSRVSMFPTRLSPCSSATEQTERDSATAYRMTVEVLGTVL
GTAIQQQIVGQADTPCFQDFNSSTVASQSANHTHGTTSHRETQKAYLLAAGVIVCIYIICAV
ILILGVREQREPYEAQQSEPIAYFRGLRLVM SHGPYIKLITGFLFTSLAFMLVEGNFVLFCT
YTLGFRNEFQNLLLAIMLSATLTIPIWQWFLTRFGKKTAVYVGISSAVPFLILVALMESNLI
ITYAVAVAAGISVAAAFLLPWSMLPDVIDDFHLKQPHFHGTEPIFFSFYVFFTKFASGVSLG
ISTLSLDFAGYQTRGCSQPERVKFTLNMLVTMAPIVLILLGLLLFKMYPIDEERRRQNKKAL
QALRDEASSSGCSETDSTELASIL

12/330

FIGURE 13

GGGAAACGCAAAAGGCATACCTGCTGGCAGCGGGGGTCATTGTCTGTATCTATATAATCTGT
GCTGTCATCCTGATCCTGGGCGTGCGGGAGCAGAGAGAACCCTATGAAGCCCAGCAGTCTGA
GCCAATCGCCTACTTCCGGGGCCTACGGCTGGTCATGAGCCACGGCCCATACATCAAACCTTA
TTACTGGCTTCCTCTTCACCTCCTTGGCTTTTCATGCTGGTGGAGGGGAACTTTGTCTTGTTT
TGCACCTACACCTTGGGCTTCCGCAATGAATTCCAGAACTACTCCTGGCCATCATGCTCTC
GGCCACTTTAACCATTCCCATCTGGCAGTGGTTCTTGACCCGGTTTGGCAAGAAGACAGCTG
TATATGTTGGGATCTCATCAGCAGTGCCATTTCTCATCTTGGTGGCCCTCATGGAGAGTAAC
CTCATCATTACATATGCGGTAGCTGTGGCAGCTGGCATCAGTGTGGCAGCTGCCTTCTTACT
ACCCTGGTCCATGCTGCCTGATGTCATTGACGACTTCCATCTGAAGCAGCCCCACTTCCATG
GAACCGAGCCCCAT

FIGURE 14

GGGGCTTCGGCGCCAGCGGCCAGCGCTAGTCGGTCTGGTAAGGATTTACAAAAGGTGCAGGT
ATGAGCAGGTCTGAAGACTAACATTTTGTGAAGTTGTAAAACAGAAAACCTGTTAGAAATGT
GGTGGTTTCAGCAAGGCCTCAGTTTCCTTCCTTCAGCCCTTGTAATTTGGACATCTGCTGCT
TTCATATTTTCATACATTACTGCAGTAACACTCCACCATATAGACCCGGCTTTACCTTATAT
CAGTGACACTGGTACAGTAGCTCCAGAAAAATGCTTATTTGGGGCAATGCTAAATATTGCGG
CAGTTTTATGCATTGCTACCATTTATGTTTCGTTATAAGCAAGTTCATGCTCTGAGTCCTGAA
GAGAACGTTATCATCAAATTAACAAGGCTGGCCTTGTTACTTGGAATACTGAGTTGTTTAGG
ACTTTCTATTGTGGCAAACCTTCCAGAAAACAACCCTTTTTGCTGCACATGTAAGTGGAGCTG
TGCTTACCTTTGGTATGGGCTCATTATATATGTTTGTTTCAGACCATCCTTTCCCTACCAAATG
CAGCCCCAAAATCCATGGCAAACAAGTCTTCTGGATCAGACTGTTGTTGGTTATCTGGTGTGG
AGTAAGTGCACCTTAGCATGCTGACTTGCTCATCAGTTTTGCACAGTGGCAATTTTGGGACTG
ATTTAGAACAGAACTCCATTGGAACCCCGAGGACAAAGGTTATGTGCTTCACATGATCACT
ACTGCAGCAGAATGGTCTATGTCAATTTTCCTTCTTTGGTTTTTTTCCTGACTTACATTCGTGA
TTTTTCAGAAAATTTCTTTACGGGTGGAAGCCAATTTACATGGATTAACCCTCTATGACACTG
CACCTTGCCCTATTAACAATGAACGAACACGGCTACTTTCCAGAGATATTTTGATGAAAGGAT
AAAATATTTCTGTAATGATTATGATTCTCAGGGATTGGGGAAAGGTTACAGAAGTTGCTTA
TTCTTCTCTGAAATTTTCAACCACTTAATCAAGGCTGACAGTAACACTGATGAATGCTGATA
ATCAGGAAACATGAAAGAAGCCATTTGATAGATTATTCTAAAGGATATCATCAAGAAGACTA
TTAAAAACACCTATGCCTATACTTTTTTATCTCAGAAAATAAAGTCAAAAGACTATG

FIGURE 15

MWWFQQGLSFLPSALVIWTSAAFIIFS YITAVTLHHIDPALPYISDTGTVAPEKCLFGAMLNI
AAVLCIATIYVRYKQVHALSPEENVIIKLNKAGLVLGILSCLGLSIVANFQKTTLFAAHVSG
AVLTFGMGSLYMFVQTILSYQMOPKIHGKQVFWIRLLLVIWCGVSALSMLTCSSVLHSGNFG
TDLEQKLHWNPEDKGYVLHMITTAAEWSMSFSFFGFFLT YIRDFQKISLRVEANLHGLTLYD
TAPCPINNERTRLLSRDI

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 16

CGGACGCTTGGGCNCGCCAGCGGCCAGCGCTAGTCGGTCTGGTAAGTGCCTGATGCCGAGT
TCCGTCTCTCGGGTCTTTTCCTGGTCCCAGGCAAAGCGGAGCGGAGATCCTCAAACGGCCTA
GTGCTTCGCGCTTCCGGAGAAAATCAGCGGTCTAATTAATTCCTCTGGTTTGTGGAAGCAGT
TACCAAGAATCTTCAACCCTTTCCACAAAAGCTAATTGAGTACACGTTCTGTGAGTACA
CGTTCCTGTTGATTTACAAAAGGTGCAGGTATGAGCAGGTCTGAAGACTAACATTTTGTGAA
GTTGTAAACAGAAAACCTGTTAGAAATGTGGTGGTTTCAGCAAGGCCTCAGTTTCCTTCCT
TCAGCCCTTGTAATTTGGACATCTGCTGCTTTCATATTTTCATACATTACTGCAGTAACACT
CCACCATATAGACCCGGCTTTACCTTATATCAGTGACACTGGTACAGTANC

40941424344454647484950

FIGURE 17

CCCACGCGTCCGCCCCGCGCTGCGTCCCGGAGTGCAAGTGAGCTTCTCGGCTGCCCCGCGGG
CCGGGGTGCGGAGCCGAC**ATG**CGCCCCGCTTCTCGGCCTCCTTCTGGTCTTCGCCGGCTGCAC
CTTCGCCTTGTACTTGCTGTGACGCGACTGCCCCGCGGGCGGAGACTGGGCTCCACCGAGG
AGGCTGGAGGCAGGTGCTGTGGTTCCCCTCCGACCTGGCAGAGCTGCGGGAGCTCTCTGAG
GTCCTTCGAGAGTACCGGAAGGAGCACACAGGCCTACGTGTTCCCTGCTCTTCTGCGGCGCCTA
CCTCTACAAACAGGGCTTTGCCATCCCCGGCTCCAGCTTCCTGAATGTTTTAGCTGGTGCCT
TGTTTGGGCCATGGCTGGGGCTTCTGCTGTGCTGTGTGTTGACCTCGGTGGGTGCCACATGC
TGCTACCTGCTCTCCAGTATTTTTGGCAAACAGTTGGTGGTGTCTACTTTCCTGATAAAGT
GGCCCTGCTGCAGAGAAAGGTGGAGGAGAACAGAAACAGCTTGTTTTTTTTCTTATTGTTTT
TGAGACTTTTCCCCATGACACCAAAGTGGTTCTTGAACCTCTCGGCCCCAATTCTGAACATT
CCCATCGTGCAGTTCTTCTTCTCAGTTCTTATCGGTTTGATCCCATATAATTTTCATCTGTGT
GCAGACAGGGTCCATCCTGTCAACCCTAACCTCTCTGGATGCTCTTTTCTCCTGGGACACTG
TCTTTAAGCTGTTGGCCATTGCCATGGTGGCATTAAATTCCTGGAACCCTCATTAAAAAATTT
AGTCAGAAACATCTGCAATTGAATGAAACAAGTACTGCTAATCATATACACAGTAGAAAAGA
CACAT**TGA**TCTGGATTTTCTGTTTGCCACATCCCTGGACTCAGTTGCTTATTTGTGTAATGGA
TGTGGTCCTCTAAAGCCCCCTCATTGTTTTTGATTGCCTTCTATAGGTGATGTGGACACTGTG
CATCAATGTGCAGTGTCTTTTCAGAAAGGACACTCTGCTCTTGAAGGTGTATTACATCAGGT
TTTCAAACCAGCCCTGGTGTAGCAGACACTGCAACAGATGCCTCCTAGAAAATGCTGTTTGT
GGCCGGGCGCGGTGGCTCACGCTGTAATCCAGCACTTTGGGAGGCCGAGGCCGGTGATTC
ACAAGGTCAGGAGTTCAAGACCAGCCTGGCCAAGATGGTGAAATCCTGTCTCTAATAAAAAT
ACAAAAATTAGCCAGGCGTGGTGGCAGGCACCTGTAATCCAGCTACTCGGGAGGCTGAGGC
AGGAGAATTGCTTGAACCAAGGTGGCAGAGGTTGCAGTAAGCCAAGATCACACCACTGCACT
CCAGCCTGGGTGATAGAGTGAGACACTGTCTTGAC

FIGURE 18

MRPLLGLLLVFAGCTFALYLLSTRLPGRRLGSTEEAGGRSLWFPSDLAELRELVREYR
 KEHQAYVFLLFCGAYLYKQGFAIPGSSFLNVLGALFGPWLGLLLCCVLTSVGATCCYLLSS
 IFGQQLVVSYPDKVALLQRKVEENRNSLFFFLFLRLFPMTPNWFLNLSAPILNIPIVQFF
 FSVLIGLIPYNFICVQTGSILSTLTSLDALFSWDTVFKLLAIAMVALIPGTLIKKFSQKHLQ
 LNETSTANHIHSRKDT

Important features:**Signal peptide:**

amino acids 1-17

Transmembrane domains:

amino acids 101-123, 189-211

N-glycosylation sites.

amino acids 172-176, 250-254

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 240-244, 261-265

N-myristoylation site.

amino acids 13-19, 104-110, 115-121, 204-210

Amidation site.

amino acids 27-31

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 4-15

Protein splicing proteins.

amino acids 25-31

Sugar transport proteins.

amino acids 162-172

FIGURE 19

CCGAGGCGGGAGGAGCCCGAGGGGGCGCGAGCCCCGCATGAATCATTGTAGTCAATCATTTT
CCAGTTCTCAGCCGCTCAGTTGTGATCAAGGGACACGTGGTTTCCGAACTGCCAGCTCAGAA
TAGGAAAATAACTTGGGATTTTATATTGGAAGACATGGATCTTGCTGCCAACGAGATCAGCA
TTTATGACAACTTTCAGAGACTGTTGATTTGGTGAGACAGACCGGCCATCAGTGTGGCATG
TCAGAGAAGGCAATTGAAAAATTTATCAGACAGCTGCTGGAAAAGAATGAACCTCAGAGACC
CCCCCGCAGTATCCTCTCCTTATAGTTGTGTATAAGGTTCTCGCAACCTTGGGATTAATCT
TGCTCACTGCCTACTTTGTGATTCAACCTTTCAGCCCATTAGCACCTGAGCCAGTGCTTTCT
GGAGCTCACACCTGGCGCTCACTCATCCATCACATTAGGCTGATGTCTTGCCCATTGCCAA
GAAGTACATGTCAGAAAATAAGGGAGTTCCTCTGCATGGGGGTGATGAAGACAGACCCTTTC
CAGACTTTGACCCCTGGTGGACAAACGACTGTGAGCAGAATGAGTCAGAGCCCATTCTGCC
AACTGCACTGGCTGTGCCAGAAACACCTGAAGGTGATGCTCCTGGAAGACGCCCCAAGGAA
ATTTGAGAGGCTCCATCCACTGGTGATCAAGACGGGAAAGCCCCTGTTGGAGGAAGAGATTC
AGCATTTTTTGTGCCAGTACCCTGAGGCGACAGAAGGCTTCTCTGAAGGGTTTTTCGCCAAG
TGGTGGCGCTGCTTTCCTGAGCGGTGGTTCCCATTTCTTATCCATGGAGGAGACCTCTGAA
CAGATCACAAATGTTACGTGAGCTTTTTCTGTTTTTCACTCACCTGCCATTTCCAAAAGATG
CCTCTTTAAACAAGTGCTCCTTTCTTCACCCAGAACCTGTTGTGGGGAGTAAGATGCATAAG
ATGCCTGACCTATTTATCATTGGCAGCGGTGAGGCCATGTTGCAGCTCATCCCTCCCTTCCA
GTGCCGAAGACATTGTCAGTCTGTGGCCATGCCAATAGAGCCAGGGGATATCGGCTATGTCG
ACACCACCCACTGGAAGGTCTACGTTATAGCCAGAGGGGTCCAGCCTTTGGTCATCTGCGAT
GGAACCGCTTTCTCAGAACTGTAGGAAATAGAAGTGTGCACAGGAACAGCTTCCAGAGCCGA
AAACCAGGTTGAAAGGGGAAAAATAAAAACAAAACGATGAAACTGCAAAAA

MDLAANEISIIYDKLSETVDLVRQTGHQCGMSEKAIEKFIRQLLEKNEPQRPPPQYPLLIVVY
KVLATLGLILLTAYFVIQFSPPLAPEPVLSGAHTWRSLIHHIRLMSLPIAKKYMSENKGVPL
HGGDEDRPFPDFDPWWTNDCEQNESEPIPANCTGCAQKHLKVMLLEDAPRKFERLHPLVIKT
GKPLLEEEIQHFCLCQYPEATEGFSEGFFAKWWRCFPERWFPPFPYWRRLNRSQMLRELFV
FTHLPFPKDASLNKCSFLHPEPVVGSKMHKMPDLFIIGSGEAMLQLIPPFQCRHCQSVAMP
IEPGDIGYVDTTTHWKVYVIARGVQPLVICDGTAFSEL

CCACGGTGTCCGTTCTTCGCCCGGCGGCAGCTGTCCCCGAGGCGGGAGGAGCCCCGAGGGGCG
CGAGCCCCGCATGAATCATTGTAGTCAATCATTTTCCAGTTCTCAGCCGTTCAAGTTGTGATC
AAGGGACACGTGGTTTTCCGAACTGCCAGCTCAGAATAGGAAAATAACTTGGGATTTTATATT
GGAAGACATGGATCTTGCTGCCAACGAGATCAGCATTTATGACAAACTTTCAGAGACTGTTG
ATTTGGTGAGACAGACCGGCCATCAGTGTGGCATGTCTCAGAGAAGGCAATTGAAAAATTTATC
AGACAGCTGCTGGAAAAGAATGAACCTCAGAGACCCCCCCCCGCAGTATCCTCTCCTTATAGT
TGTGTATAAGGTTCTCGCAACCTTGGGATTAATCTTGCTCACTGCCTACTTTGTGATTCAAC
CTTTCAGCCCATTAGCACCTGAGCCAGTGCTTTGTGGAGCTCAC

CCACACGCGTCCGCCACACGCTCCGGCTGAACACCTCTTTGGAGTCAGCCACTGATGAGG
CAGGGTCCCCACTTGCAGCTGCAGCAGCTGCAGCAGCTGCAGAGCGCTGCTCCTGGCTGGT
CCACTGGTGCACACGCTGCTAGACCGTGCTATGAGCCGCTGGGGCTGCAGTGGGGACTGCC
CTCCCTGCCACCCACCAATGGCAGCCCCACCTTCTTTGAAGACTTCCAGGCTTTTTGTGCCA
CACCAGAATGGCGCCACTTTCATCGACAAACAGGTACAGCCAACCA**ATG**TCCCAGTTTCGAAATG
GACACGTATGCTAAGAGGCCACGACCTTATGTCAAGTTTCTGGAATGCCTGCTATGACATGCT
TATGAGCAGTGGGCAGCGGCGCCAGTGGGAGCGCGCCAGAGCTCGTGGGGCTTCCAGGAGC
TGGTGCTGGAACCTGCGCAGAGGGCGGGCGCGCCTGGAGGGGCTACGCTACACGGCAGTGCTG
AAGCAGCAGGCAACGCAGCACTCCATGGCCCTGCTGCACTGGGGGGCGCTGTGGCGCCAGCT
CGCCAGCCCATGTGGGGCCTGGGCGCTGAGGGACACTCCCATCCCCCGCTGGAACCTGTCCA
GCGCCGAGACATATTACGCATGCGTCTGAAGCTGGTGCCCAACCATCACTTCGACCCCTCAC
CTGGAAGCCAGCGCTCTCCGAGACAATCTGGGTGAGGTTCCCTTGACACCCACCGAGGAGG
CTCACTGCCTCTGGCAGTGACCAAGAGGCCAAAGTGAGCACCCACCCAGATTGCTGCAGG
AGGACCAGCTCGGCGAGGACGAGCTGGCTGAGCTGGAGACCCCGATGGAGGCAGCAGAAGTG
GATGAGCAGCTGAGAAGCTGGTGCTGTGCGCCGAGTGGCAGCTGGTGACGGTAGTGCCCGT
GGTCCCAGGGCTGCTGGAGGTCAACACAGAAATGTATACCTTACGATGGCAGCACTGAGC
GCGTGGAACCGAGGAGGGCATCGGCTATGATTTCCGGCGCCCACTGGCCAGCTGCGTGAG
GTCCACCTGCGGCGTTTCAACCTGCGCCGTTACGCACTTGAGCTCTTCTTTATCGATCAGGC
CAACTACTTCTCAACTTCCCATGCAAGGTGGGCACGACCCAGTCTCATCTCCTAGCCAGA
CTCCGAGACCCAGCCTGGCCCCATCCCAACCCATAACCGAGGTACGGAACCCAGGTGTACTCG
TGGCTCCTGCGCTACGGCCCCCTCTCAAGCTACCTAAGCAGCCGCTCCCCCAGGAGAT
GCTGCGTGCCTCAGGCCTTACCCAGAAATGGGTACAGCGTGAGATATCAACTTCGAGTACT
TGATGCAACTCAACACCATTGCGGGGCGGACCTACAATGACCTGTCTCAGTACCCTGTGTTT
CCCTGGGTCTGTCAGGACTACGTGTCCCAACCCCTGGACCTCAGCAACCCAGCCGTCTTCCG
GGACCTGTCTAAGCCCCATCGGTGTTGGTGAACCCCAAGCATGCCAGCTCGTGAGGGGAGAAGT
ATGAAAGCTTTGAGGACCCAGCAGGGACCATTGACAAGTTCCACTATGGCACCCTACTCTCC
AATGCAGCAGGCGTGATGCACTACCTCATCCGCGTGGAGCCCTTCACTTCCCTGCACGTCCA
GCTGCAAAGTGGCCGCTTTGACTGCTCCGACCGGCAGTTCCACTCGGTGGCGGCAGCCTGGC
AGGCACGCCTGGAGAGCCCTGCCGATGTGAAGGAGCTCATCCCGGAATTCTTCTACTTTCT
GACTTCTTGAGAACGACGAGTGGTTTACCTTGGCTGGCTGTCTCCAGCTGACCAACGAGAAGGT
AGGCGATGTGGTGCTACCCCCGTGGGCCAGCTCTCCTGAGGACTTCATCCAGCAGCAGCCGC
AGGCTCTGGAGTCGGAGTATGTGTCTGCACACCTACACGAGTGGATCGACCTCATCTTTGGC
TACAGCAGCAGCGGGGGCCAGCCGCCGAGGAGGCCCTCAATGTCTTCTATTACTGCACCTATGA
GGGGCTGTAGACCTGGACCATGTGACAGATGAGCGGGAACGGAAGGCTCTGGAGGGCATTAT
TCAGCAACTTTGGGACAGTCCCTGTGACGTGCTGAAGGAGCCACATCCAACCTCGGCTCTCA
GCTGAGGAAGCAGCCCATCGCCTTGACGCGCTGGACACTAAGTACCTACCTAGCATCTTCCAGCA
CCTGGACGAACTCAAGGCATTCTTCGAGAGGCTGACTGTGAGTGCCAGTGGGCTGCTGGGCA
CCCACAGCTGGTTGCCCTATGACCGCAACATAAGCAACTACTTCAGCTTCAGCAAAGACCCC
ACCATTGGGCGACCCACAAGAGCGCAGCACTGCTGAGTGGCCCGTGGGTGCCAGGCAGTGGTGT
GAGTGGACAAGCACTGGCAGTGGCCCGGATGGAAAGCTGCTATTACGCGGTGGCCACTGGG
ATGGCAGCCTGCGGGTGACTGCACCTACCCCGTGGCAAGCTGTTGAGCCAGCTCAGCTGCCAC
CTTGATGTAGTAACCTGCCTTGCACTGGACACCTGTGGCATCTACCTCATCTCAGGCTCCCG
GGACACCACGTGCATGGTGTGGCGGCTCCTGCATCAGGGTGGTCTGTGAGTAGGCCTGGCAC
CAAAGCCTGTGACGCTCCTGTATGGGCATGGGGCTGCAGTGAGCTGTGTGGCCATCAGCACT
GAACTTGACATGGCTGTCTGGATCTGAGGATGGAAGTGTGATCATACACACTGTACGCCG
CGGACAGTTTGTAGCGGCACTACGGCCTCTGGGTGCCACATTCCCTGGACCTATTTTCCACC
TGGCATTGGGGTCCGAAGGCCAGATTGTGGTACAGAGCTCAGCGTGGGAACGTCTTGGGGCC
CAGGTACACTACTCCTTGACCTGTATTGAGTCAATGGGAAGTTGCGGGCTTCACTGCCCT
GGCAGAGCAGCCTACAGCCCTGACGGTGACAGAGGACTTTGTGTTGCTGGGCACCGCCAGT
GCGCCCTGCACATCTTCCAACATAAACACACTGCTCCCGCCGCGCTTCCCTTGCCCATGAAG
GTGGCCATCCGCAGCGTGGCCGTGACCAAGGAGCGCAGCCAGTGGTGGTGGGCTGGAGGA
TGGCAAGCTCATCGTGGTGGTGCAGGGGCGACCCCTCTGAGGTGCGCAGCAGCCAGTTCGCGC
GGAAGCTGTGGCGGTCCTCGCGGCGCATCTCCAGGTGTCTCGGGAGAGACGGAATACAAC
CCTACTGAGGCGCG**CTGA**ACCTGGCCAGTCCGGCTGCTCGGGCCCCCGCCCCGGCAGGCCTG
CCCCGGGAGGCCCGCCAGAGTCCGCGGGAACACCCCGGGGTGGCGAGCCAGGGGGTGA
GCGGGGGCCACCCTGCCAGCTCAGGGATTGGCGGGCGATGTTACCCCTCAGGGATTGGCG
GGCGGAAGTCCCGCCCCCTCGCCGGCTGAGGGGGCCGCCCTGAGGGCCAGCACTGGCGTCT

FIGURE 23

MSQFEMDTYAKSHDLMSSGFWNACYDMLMSSGQRRQWERAQSRRAFQELVLEPAQRRARLEGL
RYTAVLKQQATQHSMALLHWGALWRQLASPCGAWALRDTPIPRWKLSSAETYSRMRCLKLVPN
HHFDPHLEASALRDNLGEVPLTPTEEASLPLAVTKEAKVSTPPELLQEDQLGEDELAELETP
MEAAELDEQREKLVLSAECQLVTVVAVVPGLLEVTTQNVFYFDGSTERVETEEGIGYDFRRP
LAQLREVHLRFRNLRRSALELFFIDQANYFLNFPCKVGTTPVSSPSQTPRPQPGPIPPHTQV
RNQVYSWLLRLRPPSQGYLSSRSPQEMLRASGLTQKWVQREISNFEYLMQLNTIAGRTYNDL
SQYPVFPWVLQDYVSPTLDLSNPAVFRDLSKPIGVNPKHAQLVREKYESFEDPAGTIDKFH
YGTHYSNAAGVMHYLIRVEPFTSLHVQLQSGRFDCSDRQFHSVAAAWQARLESPADVKEIIP
EFFYFPDFLENQNGFDLGCLQLTNEKVGDVVLPPWASSPEDFIQQHRQALESEYVSAHLHEW
IDLIFGYKQRGPAEEALNVFYCYTYEGAVDLDHVTDERERKALEGIISNFGQTPCQLLKEP
HPTRLSAEEAAHRLARLDTNSPSIFQHLDELKAFFAEVTVSASGLLGTHSWLPYDRNISNYF
SFSKDPTMGSHKTQRLLSGPWVPGSGVSGQALAVAPDGKLLFSGGHWGSLRVTALPRGKLL
SQLSCHLDVVTCLALDTCGIY LISGSRDTTCMVWRL LHQGGLSVGLAPKPVQVLYGHGA AVS
CVAISTELDMAVSGSEDGTVIIHTVRRGQFVAALRPLGATFPGP I FHLALGSEGQIVVQSSA
WERPGAQVTYSLHLYSVNGKLRASLPLAEQPTALTVTEDFVLLGTAQCALHILQLNTLLPAA
PPLPMKVAIRSVAVTKERSHVLVGLEDGKLIVVVAGQPSEVRSSQFARKLWRSSRRISQVSS
GETEYNPTEAR

N-glycosylation site.

amino acids 677-681

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 985-989

Tyrosine kinase phosphorylation site.

amino acids 56-65, 367-376, 543-551

N-myristoylation site.

amino acids 61-67, 436-442, 604-610, 610-616, 664-670, 691-697,
706-712, 711-717, 769-775, 785-791, 802-808, 820-826, 834-840,
873-879, 912-918, 954-960

FIGURE 24

CGGACGCGTGGGCGGACGCGTGGGGGCTGTGAGAAAGTGCCAATAAATACATCATGCAACCC
CACGGCCACCTTGTGAACTCCTCGTGCCCAGGGCTGATGTGCGTCTTCCAGGGCTACTCAT
CCAAAGGCCTAATCCAACGTTCTGTCTTCAATCTGCAAATCTATGGGGTCCTGGGGCTCTTC
TGGACCCTTAACTGGGTACTGGCCCTGGGCCAATGCGTCCTCGCTGGAGCCTTTGCCTCCTT
CTACTGGGCCTTCCACAAGCCCCAGGACATCCCTACCTTCCCCCTAATCTCTGCCTTCATCC
GCACACTCCGTTACCACACTGGGTCATTGGCATTGAGGCCCTCATCCTGACCCTTGTGCAG
ATAGCCCGGGTCATCTTGGAGTATATTGACCACAAGCTCAGAGGAGTGCAGAACCCTGTAGC
CCGCTGCATCATGTGCTGTTTCAAGTGCTGCCTCTGGTGTCTGGAAAAATTTATCAAGTTCC
TAAACCGCAATGCATACATCATGATCGCCATCTACGGGAAGAATTTCTGTGTCTCAGCCAAA
AATGCGTTCATGCTACTCATGCGAAACATTTGTCAGGGTGGTCGTCCTGGACAAAGTCACAGA
CCTGCTGCTGTTCTTTGGGAAGCTGCTGGTGGTCGGAGGCGTGGGGGTCCTGTCTTCTTTT
TTTTCTCCGGTCGCATCCCGGGGCTGGGTAAAGACTTTAAGAGCCCCCACCTCAACTATTAC
TGGCTGCCCATCATGACCTCCATCCTGGGGGCTATGTCATCGCCAGCGGCTTCTTCAGCGT
TTTCGGCATGTGTGTGGACACGCTCTTCCTCTGCTTCCTGGAAGACCTGGAGCGGAACAACG
GCTCCCTGGACCGGCCCTACTACATGTCCAAGAGCCTTCTAAAGATTCTGGGCAAGAAGAAC
GAGGCGCCCCCGGACAACAAGAAGAGGAAGAAGTGACAGCTCCGGCCCTGATCCAGGACTGC
ACCCACCCCCACCGTCCAGCCATCCAACCTCACTTCGCCTTACAGGTCTCCATTTTGTGGT
AAAAAAAGGTTTTAGGCCAGGCGCCGTGGCTCACGCCTGTAATCCAACACTTTGAGAGGCTG
AGGCGGGCGGATCACCTGAGTCAGGAGTTCGAGACCAGCCTGGCCAACATGGTGAAACCTCC
GTCTCTATTAAAAATACAAAAATTAGCCGAGAGTGGTGGCATGCACCTGTCATCCCAGCTAC
TCGGGAGGCTGAGGCAGGAGAATCGCTTGAACCCGGGAGGCAGAGGTTGCAGTGAGCCGAGA
TCGCGCCACTGCACTCCAACCTGGGTGACAGACTCTGTCTCCAAAACAAAACAAACAAACAA
AAAGATTTTATTAAAGATATTTTGTTAACTC

RTRGRTRGGCEKVPINTSCNPTAHLVNSSCPGLMCVFQGYSSKGLIQRSVFNLQIYGVGLGF
WTLNWVLALGQCVLAGAFASFYWAFHKPQDIPTFPLISAFIRTLRYHTGSLAFGALILTLVQ
IARVILEYIDHKLRGVQNPVARCIMCCFKCCLWCLEKFIKFLNRNAYIMIAIYGKNFCVSAK
NAFMLLMRNIVRVVLDKVTDLLLFFGKLLVVGGVGVLSFFFFSGRIPGLGKDFKSPHLNYY
WLPIMTSILGAYVIASGFFSVFGMCDTFLFLCFLEDLERNNGSLDRPYMYSKSLKILGKKN
EAPPDNKKRKK

FIGURE 27

MFVSDFRKEFYEVVQSQRVLLFVASDVDALCACKILQALFQCDHVQYTLVPVSGWQELETAF
LEHKEQFHYFILINCGANVDLLDILQPDEDTIFFVCDSHRPVNVVNVYNDTQIKLLIKQDDD
LEVPAYEDIFRDEEEDEEHSGNDSGSEPSEKRTRLEEEIVEQTMRRRQRREWEARRRDILF
DYEQYEHGTSSAMVMFELAWMLSKDLNDMLWVAIVGLTDQWVQDKITQMKYVTDVGVLRH
VSRHNRNEDEENTLSVDCTRISFEYDLRLVLYQHWSLHDSL CNTSYTAARFKLWSVHGQKR
LQEF LADMGLPLKQVKQKFQAMDISLKENLREMIEESANKFGMKDMRVQTF SIHFGFKHKFL
ASDVVFATMSLMESPEKDGSGTDHFIQALDSLRSRNL DKLYHGLELAKKQLRATQQTIASCL
CTNLVISQGPFLYCSLMEGTPDVMLFSRPASLSLLSKHLLKSFVCSTKNRRCKLLPLVMAAP
LSMEHGTVTTVVGIPPETDSSDRKNFFGRAFEKAAESTSSRMLHNNHFDLSVIELKAEDRSKFL
DALISLLS

FIGURE 28

GTACCTCAGCGCGAGCGCCAGGCGTCCGGCCGCCGTGGCTATGNTCGTGTCCGATTTCCGCA
AAGAGTTCTACGAGGTGGTCCAGAGCCAGAGGGTCCTTCTCTTCGTGGCCTCGGANGTGGAT
GCTCTGTGTGCGTGCAAGATCCTTCAGGCCTTGTTCCAGTGTGACCANGTGCAATATANGCT
GGTTCAGTTTCTGGGTGGCAAGAACTTGAACTGCATTTCTTGAGCATAAAGAACAGTTTC
ATTATTTTATTCTCATAAACTGTGGAGCTAATGTAGACCTATTGGATATTCTTCAACCTGAT
GAAGACACTATATTCTTTGTGTGTGACACCCATAGGCCAGTCAATGTTGTCAATGTATACAA
CGATACCC

CAGGAACCCCTCTCTTTGGGTCTGGATTGGGACCCCTTTCCAGTACCATTTTTTCTAGTGAAC
CACGAAGGGACGATACCAGAAAACACCCTCAACCCAAAGGAAATAGACTACAGCCCCAATTG
GCTGACTTTTGGCTATAGAAAAAAGAAAGGAACGAAAAGAGACAGTTTTTTTTTGGAAAGCTAA
GTCTTCCCTTTATCGAGTCAAGAAACCCCCCTTCTTGAGCTATTTACAGCTTTTAAACAATT
GAGTAAAGTACGCTCCGGTCACCATGTCAGACGCCGCCCTGGGTCCCCTCTGGGCAGCGCTC
CTGCTCTTTCTCCTGATGTGTGAGATCCGTATGGTGGAGCTCACCTTTGACAGAGCTGTGGC
CAGCGGCTGCCAACGGTCTGTGACTCTGAGGACCCCTGGATCCTGCCCATGTATCCTCAG
CCTCTTCCCTCCGGCCGCCCCACGCCCTGCCTGAGATCAGACCCTACATTAATATCACCATC
CTGAAGGTTGACAAAGGGGACCCAGGGCCCAATGGGGCTGCCAGGGTACATGGGCAGGGAGGG
TCCCCAAGGGGAGCCTGGCCCTCAGGGCAGCAAGGGTGACAAGGGGAGATGGGCAGCCCGG
GCGCCCCGTGCCAGAAGCGCTTCTTCGCTTCTCAGTGGGGCCGAAAGCGCCCTGCACAGC
GGCGAGGACTTCCAGACGCTGCTCTTCGAAAGGGTCTTTGTGAACCTTGATGGGTGCTTTGA
CATGGCGACCGGCCAGTTTGTGCTCCCCTGCGTGGCATCTACTTCTTCAGCCTCAATGTGC
ACAGCTGGAATTACAAGGAGACGTACGTGCATATTATGCATAACCAGAAAGAGGCTGTCTATC
CTGTACGCGCAGCCGACGCGAGCGACGATCATGATCAGAGCCAGAGTGTGATGCTGGACCTGGC
CTACGGGGACCGCGTCTGGGTGCGGCTCTTCAAGCGCCAGCGCAGAAAGCCCATCTACAGCA
ACGACTTCGACACCTACATCACCTTCAGCGGCCACCTCATCAAGGCCGAGGACGACTGAGGG
CCTCTGGGCCACCCTCCCGGCTGGAGAGCTCAGGTGCTGGTCCCGTCCCCTGCAGGGCTCAG
TTTGCACTGCTGTGAAGCAGGAAGGCCAGGGAGGTCCCCGGGGACCTGGCATTCTGGGGAGA
CCTGTCTTATCTTTGGCTGCCATCATCCCTCCAGCCTATTTCTGCTCCTCTCTTCTCTCT
TGGACCTATTTTAAAGAAGCTTGCTAACCTAAATATTCTAGAACTTTCCAGCCTCGTAGCCC
AGCACTTCTCAAACCTTGAAATGCATGCGAATCACCCGGGGTTCGTGTTAAATGCAGATTCT
GACTCAGCAGGTCTGAGTGGGTCCAGGATTCTGTGTTTCTCATATGTTCCCTGGGTGATGCTG
ATGGGGTCAGTCTATGAACACACATGGAGCAACCAGGTCTAGGACTTTCTCAATATTCTAG
TACTTTCTGAACATTCTGGAATCCTCCCACATTTCTAGAATTTCCCAACATTTTTTTTTTCT
TGAGACAGAGTCTTGCTCTGTTGCCCAGGCTAGAGTGCAAGTGGTGCAATCTCAGTTCACTGC
AACCTCTGCCTCCCGGGTTCAAGCGATTCTTCTGCCTCAGCCTCCCTAGTGGCTGGGATTAC
AGGCGCCTGCTACCATGCCTGGCTAATTTTTGTATTTTTAGTAGAGATGGGGTTTTACCATA
TTGGCCAGGCTGGTCTTGAACCTCTGACTTCAGGTGACCCACCCGCCTCGGCCCTCTCAAAAT
GCTGGGATTACAGGTGTGAGCCACCGTGCCTGGCCAATTCCAACATTTCTAAATTTCTCAT
CCCTCCAGGGCTCCCCGTGCTATGTTCTCTTTACCCCTTCCCCCTCTCTCTGTCTCAGGCC
TGCACCACTGCAGCCACCGTTCATTTTATTCAATTAACACTGAGCACTCACTCTGTGCT
GGGTCCCGGGAAGGGTGAGGGGGTCAGACACAGGCCCTGCCCTGCCCTCAGTGACTGGCCA
GTCACGCCCAGGCGGGGAGAGATGTGTACATAGGTTTTAAAGCAGACCCAGGCTCATGGGG
GCCTGTGTTCTGGGTGTTTCAGGTGCTGCTGGTCTCCATTACCCACTGCTCCCCAAGGCTGG
TGGGACGGGGTCCCGGTGGCAGGGGCAGGTATCTCCTTCCCGTTCCTCATCCACCTGCCAG
TGCTCATCGTTACAGCAAACCCAGGGGGCCTTGGCCAGGTCAAGGGTTCTGTGAGGAGAGG
ACCCAGGAGTGTGGGGGCATTTGGGGGGTGAAGTGGCCCCCGAAGAATGGAACCCACACCCA
TAGCTCTCCCAAGCTGATACGGCATCTCGGAGAAGACCTGCCCTCTCACTGGGATCCC
CTTCTCTGCCCTCCTCCCAGGGCTCTGCCAGGGCCTTGCTCAGTCCCTTCCACCAAAGTCATCT
GAACTTCCGTTTTCCCAGGGCCTCCAGCTGCCCTCAGACACTGATGTCTGTCCCCAGGTGCT
CTCTGCCCCCTCATGCCCCCTCTCACCGGCCAGTGCCCCGACTCTCCAGGCTTTATCAAGGTG
CTAAGGCTCGGGTGGGCAGCTCCTCGTCTCAGAGCCCTCTCCGGCCTGGTGCTGCCTTTAC
AAACACCTGCAGGAGAAGGGGCCACGGAAGCCAGGCTTTAGAGCCCTCAGCAGGTCTGGG
AGCTAGAGCAAAGGAGGGACCTCAGGCCTTCCGTTTCTTCTTCCAGGGTGGGGTGGCCTGGT
GTTCCCCTAGCCTTCCAAACCCAGGTGGCCTGCCCTTCTCCCCAGAGGGAGGCGGCCTCCGC
CCATTGGTGCTCATGAGACTCTGGGGCTGAGGTGCCCCGGGGGGTGATCTCTGGTGCTCAC
AGCCGAGGGAGCCGTGGCTCCATGGCCAGATGACGGCAACAGGGTCTGACCAAGTGCCAGGA
AGACCTGTGCTATAAAACACCCTGCCTGATCCTGCCCTGCCTGACCCGCCACGCCCTGCC
GTCCAGCATGATTAAAGAATGCTGTCTCCTCTTGGAACAAAAAAAAAAAAAAAAA

FIGURE 30

MVTAALGPVWAALLLFLLMCEIRMVELTFDRAVASGCQRCCDSEDPLDPAHVSSASSSSGRPH
ALPEIRPYINITILKGDKGDPGPMGLPGYMGREGPGQGEPPQGSKGDKGEMGSPGAPCQKRF
FAFSVGRKTALHSGEDFQTLLFERVFVNLDGCFDMATGQFAAPLRGIYFFSLNVHSWNYKET
YVHIMHNQKEAVILYAQPSESRIMQSQSVMLDLAYGDRVWVRLFKRQRENAIYSNDFDTYIT
FSGHLIKAEDD

Important features:

Signal peptide:

amino acids 1-20

N-glycosylation site.

amino acids 72-75

Clq domain proteins.

amino acids 144-178, 78-111 and 84-117

FIGURE 31

ACTCGAACGCAGTTGCTTCGGGACCCAGGACCCCCTCGGGCCCGACCCGCCAGGAAAGACTG
 AGGCCGCGGCCTGCCCCGCCCCGGCTCCCTGCGCCGCCGCCCTCCCGGGACAGAAG**ATG**TG
 CTCCAGGGTCCCTCTGCTGCTGCCGCTGCTCCTGCTACTGGCCCTGGGGCCTGGGGTGCAGG
 GCTGCCCATCCGGCTGCCAGTGCAGCCAGCCACAGACAGTCTTCTGCACTGCCCGCCAGGGG
 ACCACGGTGCCCCGAGACGTGCCACCCGACACGGTGGGGCTGTACGTCTTTGAGAACGGCAT
 CACCATGCTCGACGCAGGCAGCTTTGCCGGCCTGCCGGGCCTGCAGCTCCTGGACCTGTCAC
 AGAACCAGATCGCCAGCCTGCCAGCGGGGTCTTCCAGCCACTCGCCAACCTCAGCAACCTG
 GACCTGACGGCCAACAGGCTGCATGAAATCACCAATGAGACCTTCCGTGGCCTGCGGCGCCT
 CGAGCGCCTCTACCTGGGCAAGAACCGCATCCGCCACATCCAGCCTGGTGCCTTCGACACGC
 TCGACCGCCTCCTGGAGCTCAAGCTGCAGGACAACGAGCTGCGGGCACTGCCCCCGCTGCGC
 CTGCCCCGCCTGCTGCTGCTGGACCTCAGCCACAACAGCCTCCTGGCCCTGGAGCCCGGCAT
 CCTGGACACTGCCAACGTGGAGGCGCTGCGGCTGGCTGGTCTGGGGCTGCAGCAGCTGGACG
 AGGGGCTCTTCAGCCGCTTGCGCAACCTCCACGACCTGGATGTGTCCGACAACCAGCTGGAG
 CGAGTGCCACCTGTGATCCGAGGCCTCCGGGGCCTGACGCGCCTGCGGCTGGCCGGCAACAC
 CCGCATTGCCCAGCTGCGGGCCGAGGACCTGGCCGGCCTGGCTGCCCTGCAGGAGCTGGATG
 TGAGCAACCTAAGCCTGCAGGCCCTGCCTGGCGACCTCTCGGGCCTCTTCCCCCGCCTGCGG
 CTGCTGGCAGCTGCCCCGAACCCCTTCAACTGCGTGTGCCCCCTGAGCTGGTTTGGCCCCCTG
 GGTGCGCGAGAGCCACGTCACTTGGCCAGCCCTGAGGAGACGCGCTGCCACTTCCCGCCCA
 AGAACGCTGGCCGGCTGCTCCTGGAGCTTGACTACGCCGACTTTGGCTGCCAGCCACCACC
 ACCACAGCCACAGTGCCACCACGAGGCCCGTGGTGCGGGAGCCCACAGCCTTGTCTTCTAG
 CTTGGCTCCTACCTGGCTTAGCCCCACAGCGCCGGCCACTGAGGCCCCCAGCCCGCCCTCCA
 CTGCCCCACCGACTGTAGGGCCTGTCCCCCAGCCCCAGGACTGCCACCGTCCACCTGCCTC
 AATGGGGGCACATGCCACCTGGGGACACGGCACCACTGGCGTGCTTGTGCCCCGAAGGCTT
 CACGGGCCTGTACTGTGAGAGCCAGATGGGGCAGGGGACACGGCCAGCCCTACACCAGTCA
 CGCCGAGGGCCACCACGGTCCCTGACCCTGGGCATCGAGCCGGTGAGCCCCACCTCCCTGCGC
 GTGGGGCTGCAGCGCTACCTCCAGGGGAGCTCCGTGCAGCTCAGGAGCCTCCGTCTCACCTA
 TCGCAACCTATCGGGCCCTGATAAGCGGCTGGTGACGCTGCGACTGCCTGCCTCGCTCGCTG
 AGTACACGGTCACCCAGCTGCGGCCCAACGCCACTTACTCCGTCTGTGTGCATGCCTTTGGGG
 CCCGGGCGGGTGCCGGAGGGCGAGGAGGCCTGCGGGGAGGGCCATACACCCCCAGCCGTCCA
 CTCCAACCACGCCCCAGTCACCCAGGCCCGCGAGGGGAACCTGCCGCTCCTCATTGCGCCCG
 CCTGGCCGCGGTGCTCCTGGCCGCGCTGGCTGCGGTGGGGGCAGCCTACTGTGTGCGGCGG
 GGGCGGGCCATGGCAGCAGCGGCTCAGGACAAAGGGCAGGTGGGGCCAGGGGCTGGGCCCT
 GGAAGTGGAGGGAGTGAAGTCCCCCTTGGAGCCAGGCCCGAAGGCAACAGAGGGCGGTGGAG
 AGGCCCTGCCCAGCGGGTCTGAGTGTGAGGTGCCACTCATGGGCTTCCAGGGCCTGGCCTC
 CAGTCACCCCTCCACGCAAAGCCCTACAT**CTAA**GCCAGAGAGAGACAGGGCAGCTGGGGCCG
 GGCTCTCAGCCAGTGAGATGGCCAGCCCCCTCCTGCTGCCACACCACGTAAGTTCTCAGTCC
 CAACCTCGGGGATGTGTGCAGACAGGGCTGTGTGACCACAGCTGGGCCCTGTTCCCTCTGGA
 CCTCGGTCTCCTCATCTGTGAGATGCTGTGGCCAGCTGACGAGCCCTAACGTCCCCAGAAC
 CGAGTGCCATGAGGACAGTGTCCGCCCTGCCCTCCGCAACGTGCAGTCCCTGGGCACGGCG
 GGCCCTGCCATGTGCTGGTAACGCATGCCTGGGTCCCTGCTGGGCTCTCCCACTCCAGGCGGA
 CCCTGGGGGCCAGTGAAGGAAGCTCCCGGAAAGAGCAGAGGGAGAGCGGGTAGGCGGCTGTG
 TGAATCTAGTCTTGGCCCCAGGAAGCGAAGGAACAAAAGAACTGGAAAGGAAGATGCTTTA
 GGAACATGTTTTGCTTTTTTAAATATATATATTTATAAGAGATCCTTTCCCATTTATTCTG
 GGAAGATGTTTTTCAAATCAGAGACAAGGACTTTGGTTTTTTGTAAGACAAACGATGATATG
 AAGGCCTTTTGTAAAGAAAAAATAAAAGATGAAGTGTGAAA

FIGURE 32

MCSRVPLLLLPLLLLLLALGPGVQGCPSGCQCSQPQTVFCTARQGTTPRDVPPDTVGLYVFEN
GITMLDAGSFAGLPGLQLLDLSQNQIASLPSGVFQPLANLSNLDLTANRLHEITNETFRGLR
RLERLYLGKNRIRHIQPGAFDTLDRLLELKLQDNELRALPPLRLPRLLLLDLSHNSLLALEP
GILDTANVEALRLAGLGLQQLDDEGLFSRLRNLDLDVSDNQLERVPPVIRGLRGLTRLRLAG
NTRIAQLRPEDLAGLAALQELDVSNSLSLQALPGDLSGLFPRLRLLAAARNPFNCVCPLSWFG
PWVRESHVTLASPEETRCHFPKPNAGRLLLELDYADFGCPATTTTATVPTRPVVREPTALS
SSLAPTWLSPTAPATEAPSPSTAPPTVGPVPQPQDCPPSTCLNGGTCHLGTRHHLACLCPE
GFTGLYCESQMGQGTRPSPTPVTPRPPRSLTLGIEPVSPTSRLRVGLQRYLQGSSVQLRSLRL
TYRNLSGPDKRLVTLRLPASLAEYTVTQLRPNATYSVCMPLGPGRVPEGEEACGEAHTPPA
VHSNHAPVTQAREGNLPLLIAPALAAVLLAALAAVGAAYCVRRGRAMAAAAQDKGQVGPGAG
PLELEGVKVPLEPGPKATEGGGEALPSGSECEVPLMGFPGLQSPHAKPYI

GAATCATCCACGCACCTGCAGCTCTGCTGAGAGAGTGC AAGCCGTGGGGTTTTGAGCTCAT
CTTCATCATTTCATATGAGGAAATAAGTGGTAAAATCCTTGGAAATACA**ATG**AGACTCATCAG
AAACATTTACATATTTTGTAGTATTGTTATGACAGCAGAGGGTGATGCTCCAGAGCTGCCAG
AAGAAAGGGAACTGATGACCAACTGCTCCAACATGTCTCTAAGAAAGGTTCCCGCAGACTTG
ACCCAGGCCACAACGACACTGGATTTATCCTATAACCTCCTTTTTCAACTCCAGAGTTTCTGAG
TTTTCAATTCTGTCTCCAACTGAGAGTTTTGATTCTATGCCATAACAGAATTCAACAGCTGG
ATCTCAAAACCTTTGAATTCAACAAGGAGTTAAGATATTTAGATTTTGTCTAATAACAGACTG
AAGAGTGTAACCTTGGTATTTACTGGCAGGTCTCAGGTATTTAGATCTTTCTTTTAATGACTT
TGACCCATGCCTATCTGTGAGGAAGCTGGCAACATGTACACACCTGGAAATCCTAGGTTTGA
TGGGGCAAAAATAACAAAATCAGATTTCCAGAAAATTTGCTCATCTGCATCTAAATACTGT
TTCTTAGGATTCAGAACTCTTCTCATTTAAGAAGGTAGCCTGCCACTTAAACACAAC
AAAAGTGCACATTTGTTTTACCAATGGACACAAATTTCTGGGTTCTTTTGCCTGATGGAATCA
AGACTTCAAAAATATTAGAAATGACAAATATAGATGGCAAAAGCCAATTTGTAAGTTATGAA
ATGCAACGAAATCTTAGTTTAGAAAATGCTAAGACATCGGTTCTATTGCTTAATAAAGTTGA
TTTACTCTGGGACGACCTTTTCTTATCTTACAATTTGTTTGGCATACATCAGTGGAACT
TTCAGATCCGAAATGTGACTTTTGGTGGTAAGCTTATCTTGACCACAATTCATTTGACTAC
TCAAATACTGTAATGAGAACTATAAAAATTGGAGCATGTACATTTAGAGTGTTTTACATTCA
ACAGGATAAAATCTATTTGCTTTTGACCAAAATGGACATAGAAAACCTGACAATATCAAATG
CACAATATGCCACACATGCTTTTCCCGAATTATCCTACGAAATCCAATATTTAAATTTTGCC
AATAATATCTTAACAGACGAGTTGTTTAAAGAACTATCCAAGTGCCTCAGCTGAAAACCTCT
CATTTTGAATGGCAATAAACTGGAGACACTTTCTTTAGTAAGTTGCTTGTCTAACACACAC
CCTTGGAACTTGGATCTGAGTCAAAATCTATTACAACATAAAAATGATGAAAATTTGCTCA
TGGCCAGAACTGTGGTCAATATGAATCTGTACATAATAAATTGTCTGATTCTGTCTTCAG
TGCTCTTGCCCAAAAGTATTCAAATACTTGACCTAAATAATAACCAATCCAACCTGTACCTA
AAGAGACTATTCTATCTGATGGCCTTACGAGAATAAATGATGATTTAATTTCTTAATCTGAT
CTCCCTGGATGCAGTCAATTTAGTAGACTTTTCACTTCTGAACTTGAATGAATCTTATCT
CAGCCCATCTCTGGATTTTGTTCAGAGCTGCCAGGAAGTTAAAACTCTAAATGCGGGAAGAA
ATCCATTTCCGGTGTACCTGTGAATTAATAAATTTTCACTCAGCTTGAACATATTCAGAGGTC
ATGATGGTTGGATGGTCAAGTTATACACCTGTGAATACCTTTAAACCTAAGGGGAACTAG
GTTAAAAGACGTTTCTATCTCCACGAATTAATCTTGAACACAGCTCTGTTGATTGTACCATTTG
TGGTTATTATGCTAGTTCTGGGGTTGGCTGTGGCCTTCTGCTGTCTCCACTTTGATCTGCC
TGGTATCTCAGGATGCTAGGTCAATGCACACAAACATGGCACAGGGTTAGGAAAAACAACCCA
AGAACAACCTCAAGAGAAATGTCCGATTCACGCATTTATTTTATACAGTGAACATGATTCTC
TGTGGGTGAAGAAATGAATTCATCCCCAATCTAGAGAAGGAAGATGGTTCTATCTTGATTTGC
CTTTATGAAAGCTACTTTGACCTGGCAAAAGCATTAGTGAATAATTGTAAGCTTCATTGA
GAAAAGCTATAAGTCCATCTTTGTTTTGTCTCCCACTTTGTCCAGAATGAGTGGTGCCATT
ATGAATTTCTACTTTGCCCACCACAATCTCTTCCATGAAAATTTCTGATCATATAATTTCTTATC
TTACTGGAAACCCATCTCATTTGCTATTTCCACCCAGGTATCATAACTGAAAGCTCTCCT
GAAAAAAAAGCATACTTGAATGGCCCAAGGATAGGCGTAAATGTGGGCTTTTCTGGGCAA
ACCTTCGAGCTGCTATTAATGTTAATGTATTAGCCACGAGAAATGTATGAAGTGCAGACA
TTCACAGAGTTAAATGAAGAGTCTCGAGGTTCTACAATCTCTCTGATGAGAACAGATTTGTCT
ATAAAATCCACAGTCTTTGGGAAGTTGGGGACCACATACACTGTTGGGATGTACATTGATA
CAACCTTTTATGATGGCAATTTGACAATATTTATAAATAAAAAATGGTTATTCCCTTCATA
TCAGTTTCTAGAAGGATTTCTAAGAAATGTATCCTATAGAAAACACCTTCACAAGTTTATAAGG
GCTTATGGAAAAAGGTGTTTCATCCCAGGATTTGTTTATAATCATGAAAAATGTGGCCAGGTGC
AGTGGCTCACTCTTGTAAATCCAGCACTATGGGAGGCCAAGGTGGGTGACCACGAGGTCAA
GAGATGGAGACCATCCTGGCCAACATGGTGAAACCTGTCTCTACTAAAAATACAAAAATTA
GCTGGGCGTGATGGTGCACGCCTGTAGTCCCAGCTACTTTGGGAGGCTGAGGCAGGAGAATCG
CTTGAACCCGGGAGGTGGCAGTTGCAGTGAGCTGAGATCGAGCCACTGCACCTCCAGCCTGGT
GACAGAGCGAGACTCCATCTCAAAAAAAGAAAAAAGAAAAAAGAAAAAATGAAAAACATCC
TCATGGCCACAAAAATAAGGTCTAATTCAATAAATTATAGTACATTAATGTAATATAATATTA
CATGCCACTAAAAAGAATAAGGTAGCTGTATATTTCTGGTATGGAAAAACATATTAATAT
GTTATAAATCTATTAGGTTGGTGCAAAACTAATTGTGGTTTTTGGCATTGAAATGGCATTGAA
ATAAAGTGTAAGAATACTATACAGATGTAGTAAACAGTGGTTGGTCTGGGAGGTTGGA
TTACAGGGAGCATTTGATTTCTATGTTGTGTATTTCTATAATGTTTGAATTTGTTAGATGA
ATCTGTATTTCTTTTATAAGTAGAAAAAATAAAGATAGTTTTTACAGCCT

FIGURE 34

MRLIRNIYIFCSIVMTAEGDAPELPEERELMTNCSNMSLRKVPADLTPATTTLDLSYNLLFQ
LQSSDFHSVSKLRVLILCHNRIQQDLKTFEENKELRYLDLSNNRLKSVTWYLLAGLRYLDL
SFNDFDTMPICEEAGNMSHLEILGLSGAKIQKSDFQKIAHLHLNTVFLGFRTLPHYEEGSLP
ILNTTKLHIVLPMDTNFWVLLRDGIKTSKILEMTNIDGKSQFVSYEMQRNLSLENAKTSVLL
LNKVDLLWDDLFLILQFVWHTSVEHFQIRNVTFGGKAYLDHNSFDYSNTVMRTIKLEHVHFR
VFYIQQDKIYLLLTkMDIENLTISNAQMPHMLFPNYPTKFQYLNFANNILTDELfKRTIQLP
HLKTLILNGNKLETLSLVSCFANNTPLEHLDLSONLLQHKNDENCswPETVVNMNLSYNKLS
DSVFRCLPKSIQILDNLNNQIQTVPKETIHLMALRELNIAFNFLTDLPGCSHFSRLSVLNIE
MNFILSPSLDFVQSCQEVKTLNAGRNPFRCTCELKNFIQLETYSEVMMVGWSDSYTCEYPLN
LRGTRLKDVHLHELSCNTALLIVTIVVIMLVGLAVAFCCCLHFDLPWYLRMLGQCTQTWHRV
RKTTQEQLKRNVRFHAFISYSEHDSLWVKNELIPNLEKEDGSILICLYESYFDPGKSISENI
VSFIEKSYKSIFVLSPNFVQNEWCHYEFYFAHHNLFHENSdHIILILLEPIPFYCIPTRYHK
LKALLEKKAYLEWPKDRRKCGLFWANLRAAINVNVLATREMYELQTFTELNEESRGSTISLM
RTDCL

FIGURE 35

[illegible]

FIGURE 36

MSRPGTATPALALVLLAVTLAGVGAQGALEDPDYYGQEIWSREPYYPARPEPELETFSPPLP
AGPGEEWERRPQEPRPPKRATKPKKAPKREKSAPEPPPPGKHSNKKVMRTKSSEKAANDDHS
VRVAREDVRESCPPLGLETLKITDFQLHASTVKRYGLGAHRGRLNIQAGINENDFYDGAWCA
GRNDLQQWIEVDARRLTRFTGVITQGRNSLWLSDWVTSYKVMVSNDSTWVTVKNGSGDMI F
EGNSEKEIPVLNELPVPVMVARYIRINPQSWFDNGSICMRMEILGCPLPDPNNYYHRRNEMTT
TDDLDFKHHNYKEMRQLMKVVNEMCPNITRIYNIGKSHQGLKLYAVEISDHPGEHEVGEPEF
HYIAGAHGNEVLGRELLLLLVQFVCQEYLARNARIVHLVEETRIHVLP SLNPDGYEKAYEGG
SELGGWSLGRWTHDGIDINNNFPDLNTLLWEAEDRQNVPRKVPNHYIAIPEWFLSENATVAA
ETRAVIAWMEKIPFVLGGNLQGGELVVAYPYDLVRS PWKTQEHTPTPDDHVFRWLAYS YAST
HRLMTDARRRVCHTEDFQKEEGTVNGASWHTVAGSLNDFS YLHTNCFELSIYVGCDKYPHES
QLPEEWENNRESLIVFMEQVHRGIKGLVRDSHGKGIPNAIISVEGINHDIRTANDGDYWRL L
NPGEYVVTAKAEGFTASTKNCMVGYDMGATRCDFTL SKTNMARIREIMEKFGKQPVSLPARR
LKLGRKRQRG

FIGURE 37

CTAAGAGGACAAGATGAGGCCCGGCTCTCATTTCTCCTAGCCCTTCTGTTCTTCCTTGGCCAAGCTGCAGGGG
ATTTGGGGGATGTGGGACCTCCAATTCCCAGCCCCGGCTTCAGCTCTTTCCCAGGTGTTGACTCCAGCTCCAGC
TTCAGCTCCAGCTCCAGGTCCGGGCTCCAGCTCCAGCCGAGCTTAGGCAGCGGAGGTTCTGTGTCCCAGTTGTT
TTCCAATTTACCCGGCTCCGTGGATGACCGTGGGACCTGCCAGTGCTCTGTTTCCCTGCCAGACACCACCTTTC
CCGTGGACAGAGTGGAACGCTTGGAATTCACAGCTCATGTTCTTTCTCAGAAGTTTGAGAAAGAACTTTCTAAA
GTGAGGGAATATGTCCAATTAATTAGTGTGTATGAAAAGAACTGTAAACCTAACTGTCCGAATTGACATCAT
GGAGAAGGATACCATTTCTTACACTGAACTGGACTTCGAGCTGATCAAGGTAGAAGTGAAGGAGATGGAAAAAC
TGGTCATACAGCTGAAGGAGAGTTTTGGTGGAAGCTCAGAAATTGTTGACCAGCTGGAGGTGGAGATAAGAAAT
ATGACTCTCTTGGTAGAGAAGCTTGAGACACTAGACAAAAACAATGTCCTTGCCATTGCGCCGAGAAATCGTGGC
TCTGAAGACCAAGCTGAAAGAGTGTGAGGCCTCTAAAGATCAAAACACCCCTGTGTCACCCCTCCTCCCACTC
CAGGGAGCTGTGGTTCATGGTGGTGTGGTGAACATCAGCAAACCGTCTGTGGTTTCAGCTCAACTGGAGAGGGTTT
TCTTATCTATATGGTGCTTGGGGTAGGGATTACTCTCCCCAGCATCCAAACAAAGGACTGTATTGGGTGGCGCC
ATTGAATACAGATGGGAGACTGTTGGAGTATTATAGACTGTACAACACACTGGATGATTTGCTATTGTATATAA
ATGCTCGAGAGTTGCGGATCACCTATGGCCAAGGTAGTGGTACAGCAGTTTACAACAACAACATGTACGTCAAC
ATGTACAACACCGGGAATATTGCCAGAGTTAACCTGACCACCAACACGATTGCTGTGACTCAAACCTCTCCCTAA
TGCTGCCTATAATAACCGCTTTTCATATGCTAATGTTGCTTGGCAAGATATTGACTTTGCTGTGGATGAGAAAT
GATTGTGGGTTATTTATTCAACTGAAGCCAGCACTGGTAACATGGTGATTAGTAAACTCAATGACACCACACTT
CAGGTGCTAAACACTTGGTATACCAAGCAGTATAAACCATCTGCTTCTAACGCCTTCATGGTATGTGGGTTCT
GTATGCCACCCGTACTATGAACACCAGAACAGAAGAGATTTTTTACTATTATGACACAAACACAGGAAAGAGG
GCAAACCTAGACATTGTAATGCATAAGATGCAGGAAAAAGTGCAGAGCATTAACCTATAACCCCTTTTGACCAGAAA
CTTTATGTCTATAACGATGGTTACCTTCTGAATTATGATCTTTCTGTCTTGCAGAAGCCCCAGTAAAGCTGTTA
GGAGTTAGGGTGAAAGAGAAAATGTTTGTGAAAAAATAGTCTTCTCCACTTACTTAGATATCTGCAGGGGTGT
CTAAAAGTGTGTTCATTTTGCAGCAATGTTTAGGTGCATAGTTCTACCACACTAGAGATCTAGGACATTTGTCT
TGATTTGGTGAGTTCTCTTGGGAATCATCTGCCCTCTCAGGCGCATTTTGCAATAAAGTCTGTCTAGGGTGGGA
TTGTCAGAGGTCTAGGGGCACTGTGGGCCTAGTGAAGCCTACTGTGAGGAGGCTTCACTAGAAGCCTTAAATTA
GGAATTAAGGAACCTAAAACTCAGTATGGCGTCTAGGGATTCTTTGTACAGGAAATATTGCCCAATGACTAGTC
CTCATCCATGTAGCACCCTAATTCTTCCATGCCTGGAAGAAACCTGGGGACTTAGTTAGGTAGATTAATATCT
GGAGCTCCTCGAGGGACCAATCTCCAACCTTTTTTTTCCCTCACTAGCACCTGGAATGATGCTTTGTATGTGG
CAGATAAGTAAATTTGGCATGCTTATATATTCTACATCTGTAAAGTGTGAGTTTTATGGAGAGAGGCCTTTTT
ATGCATTAAATTTGACATGGCAAATAAATCCCAGAAGGATCTGTAGATGAGGCACCTGCTTTTTCTTTCTCTC
ATTGTCCACCTTACTAAAAGTCAGTAGAATCTTCTACCTCATAACTTCCTTCCAAAGGCAGCTCAGAAGATTAG
AACCAGACTTACTAACCAATTCCACCCCCACCAACCCCTTCTACTGCCTACTTTAAAAAATTAATAGTTTT
CTATGGAACCTGATCTAAGATTAGAAAAATTAATTTTCTTTAATTTTATTATGGACTTTTATTATGACTCTA
AGACTATAAGAAAATCTGATGGCAGTGACAAAGTGCTAGCATTATTTGTTATCTAATAAAGACCTTGGAGCATA
TGTGCAACTTATGAGTGTATCAGTTGTTGCATGTAATTTTTGCCTTTGTTAAGCCTGGAACCTGTAAGAAAAT
GAAAAATTAATTTTTTTTTCTAGGACGAGCTATAGAAAAGCTATTGAGAGTATCTAGTTAATCAGTGCAGTAGT
TGGAACCTTGCTGGTGTATGTGATGTGCTTCTGTGCTTTTGAATGACTTTATCATCTAGTCTTTGTCTATTTT
TCCTTTGATGTTCAAGTCTAGTCTATAGGATTGGCAGTTTAAATGCTTTACTCCCCCTTTTAAATAAATGAT
TAAATGTGCTTTGAAAAAATAAAAAAAAAAAAAAAAAAAAAA

MRPGLSFLLLALLFFFLGQAAGDLGDVGPPPIPSPGFSSFPGVDSSSSSFSSSSSRSGSSSSSRSLGS
GGSVSQLFSNFTGSDDRGTCQCSVSLPDTTFPVDRVERLEFTAHVLSQKFEKELSKVREYV
QLISVYEKKLLNLTVRIDIMEKDTISYTELDFELIKVEVKEMEKLVIQLKESFGGSSEIVDQ
LEVEIRNMTLLVEKLETLDKNNVLAIRREIVALKTKLKECEASKDQNTPVVHPPPTPGSCGH
GGVVNISKPSVVQLNWRGFSYLYGAWGRDYSPQHNPNGLYWVAPLNTDGRLLLEYRLYNTLD
DLLLYINARELRITYGQSGTAVYNNNNMYVNMNTGNIARVNLTNTTIAVTQTLPNAAAYNNR
FSYANVAWQDIDFAVDENGLWVIYSTEASTGNMVISKLNDDTLQVLNTWYTKQYKPSASNAF
MVCGLVYATRMTNTRTEEIFYYYDTNTGKEGKLDIVMHKMQEKVQSINYNPFDQKLYVYNDG
YLLNYDLSVLQKPQ

FIGURE 39

GCTCTGAAGACCAAGCTGAAAGAGTGTGAGGCCTCTAAAGATCAAACACCCCTGTCGTCCAC
CCTCCTCCCACTCCAGGGAGCTGTGGTCATGGTGGTGTGGTGAACATCAGCAAACCGTCTGT
GGTTCAGCTCAACTGGAGAGGGTTTTCTTATCTATATGGTGCTTGGGGTAGGGATTACTCTC
CCCAGCATCCAAACAAAGGNATGTATTGGGNGGCGCCATTGAATACAGATGGGAGACTGTTG
GAGTATTATAGACTGTACAACCCACTGGATGATTTGCTATTGTATATAAATGCTCGAGAGTT
GCGGATCACCTATGGCCAAGGTAGTGGTACAGCAGTTTACAACAACAACATGTACGTCAACA
TGTAACAACCCGGGNATATTGCCAGAGTTAACCTGACC

TCTCGCAGATAGTAATAATATCTCGGAAAGGCGAGAAAGAAGCTGTCTCCATCTTGTCTGTAT
CCGCTGCTCTTGTGACGTTGTGGAGATGCGGGGAGCGTCTCTGGGCTGTGCTCCATGGCGAGCT
GGATACCATGTTTGTGTGGAAGTGCCCGTGTTTGCTATGCCGATGCTGTCTAGTGGAAAC
AACTCCACTGTAACTAGATTGATCTATGCACCTTTTCTTGCTTGTGGAGTATGTGTAGCTTG
TGTAATGTTGATACCGAAGTGAAGAACAACCTGAATAAGATTCTGGATTTTGTGAGAATG
AGAAAGGTGTTGTCCTTGTAAACATTTTGGTTGGCTATAAAGCTGTATATCGTTTGTGCTTT
GGTTTGGCTATGTTCTATCTTCTTCTCTTTACTAATGATCAAAGTGAAGACTAGCAGTGA
TCCTAGAGCTGCAGTGCACAATGGATTTTGGTTCTTTAAATTTGCTGCAGCAATTGCAATTA
TTATTGGGGCATTTCTCATTCAGGAAGGAACCTTTTACAACCTGTGTGGTTTTATGTAGGCAT
GCAGGTGCCTTTTGGTTTCATCCTCATACAACCTAGTCTTACTTATTGATTGTTGCACATTCATG
GAATGAATCGTGGGTTGAAAAAATGGAAGAAGGGAACCTCGAGATGTTGGTATGCAGCCTTGT
TATCAGCTACAGCTCTGAATTATCTGCTGTCTTTAGTTGCTATCGTCCTGTTCTTTGTCTAC
TACACTCATCCAGCCAGTTGTTTCAGAAAAAAGGCGTTTCATCAGTGTCAACATGCTCCTCTG
CGTTGGTGCTTCTGTAATGTCTATCTGCAAAAATCCAAGAATCACAACCAAGATCTGGTT
TGTTACAGTCTTTCAGTAATTACAGTCTACACAATGTATTTGACATGGTCAGCTATGACCAAT
GAACCAGAAACAAATTGCAACCCAAGTCTACTAAGCATAAATTGGCTACAATACAACAAGCAC
TGTCCCAAGGAAGGGGAGTGCAGTCCAGTGGTGGCATGCTCAAGGAATTTATAGGACTAATTC
TCTTTTGTGTGTGTTTATTTCCAGCATCCGTACTTCAACAATAGTCAGGTTAATAAAA
CTGACTCTAAACAAGTGATGAATCTACATTAATGAAGATGGTGGAGCTAGAAGTGATGGATC
ACTGGAGGATGGGGACGATGTTCAACGAGCTGTAGATAATGAAAGGGATGGTGTCACTTACA
GTTATTCCTTCTTTCACTTCATGCTTTTCTGGCTTCACCTTTATATCATGATGACCCTTACC
AACTGGTCCAGGATGAACCCCTCTCGTGAGATGAAAGTCAAGTGGACAGCTGTCTGGGTGAA
AATCTCTCCAGTTGGATTGGCATCGTGTGATGTTTGGACACTCGTGGCACCACCTTGTTC
TTACAAATCGTGATTTTGGACTGAGTGTGAGACTTCTAGCATGAAAGTCCCACTTTGATTATTGC
TTATTTGAAAACAGTATTCCCAACTTTTGTAAAGTTGTGTATGTTTTGCTTCCCATGTAAC
TTCTCCAGTGTTCTGGCATGAATTAGATTTTACTGCTTGTCAATTTGTTATTTTCTTACCAA
GTGCATTGATATGTGAAGTAGAATGAATTCAGAGGAAAGTTTTATGAATATGGTGATGAGT
TAGTAAAAGTGGCCATTATTGGGCTTATTCTCTGCTCTATAGTTGTGAAATGAAGAGTAAAA
ACAAATTTGTTTGAATATTTTAAAATTATATTAGACCTTAAAGCTGTTTTAGCAAGCATTA
GCAATGTATGGCTGCCTTTTGAATATTTGATGTGTTGCCTGGCAGGATACTGCAAAGAAC
ATGGTTTTATTTTAAAATTTATAAAACAAGTCACTTAAATGCCAGTTGTCTGAAAAATCTTATA
AGGTTTTACCCTTGATACGGAATTTACACAGGTAGGAGTGTTTAGTGGACAATAGTGTAGG
TTATGGATGGAGGTGTGCGTACTAAATGAATAACGAGTAAATAATCTTACTTGGGTAGAGA
TGGCCTTTGCCAACAAAGTGAAGTGTTTTGGTTGTTTTAAACTCATGAAGTATGGGTTCAAGT
GGAAATGTTTGAAGTCTGAAGGATTTAGACAAGGTTTTGAAAAGGATAATCATGGGTTAGA
AGGAAGTGTTTGAAAGTCACTTTGAAAGTTAGTTTTGGGCCAGCAGGTGAGTCAACCTT
GGTAATCCAGCACTTTGGGAGCTTAAAGTGGGTAGATTACTTGAGCCAGGAATTCAGACCA
GCTTGGCACATGGTGAACCTGTTCTATAAAAAATAATCTGGCTTTGAGCATATGCCTGTGGTC
CAGCACTGAGAGGCTAGTGAAGATTGCTGAGCCAGAGCCAAAGGTTGCAGTGAAGCAAGTCA
CGTCACTGCACCTAGTGGCACAGAGTAAGCCAAAAAATATATATATTGAAATCAAGG
AGGCAAAATTTGACAGGGAAGGAAGTAACTGCAAAACCACTAGGCTTTAGTAGGTACTTAT
ATAAAATCTAGTCCAGTTCTCTCATTTAAAAAATAAGACACTGAAATACAGACTTAAATA
GCTCAGATAGCTAATTAGGAAATTTCAAGTTGGCCAATAATAGCATTCTCTCTGACATTTAA
AAATAATTTCTATTCAAATACATGCATATTGATTTACACCTCATACTGTGATAATTAATGT
GATGTGGATTGCTGTTGTCAGCATGACCCATAAACAGGTGAGAAGATGATGGAATGTTTT
AGAATAAATCTGCTTATAGTATACTACACAGTTCAAAAGATGTTTTAAATGCTTTTGTAT
TTACTGCCATGTAATTGAAATATATAGATTATTGTAACCTTTCAACCTGAAAATCAAGCAGT
ATGAGAGTTTAGTTATTTGTATGTGTCACTAGTGTCTAATGAAGCTTTTAAATCTACAATT
TCTTCTTTAAAAATATTTAATGTGAATGGAATATAACAATTCAGCTTAATTCCCAACC
TTATTCTGTGTGATAGCATTTGATTTCCACAATTTTGAATGGCTGTGTTTTACCTCTAAATAA
ATGAATTCAGAGAAAAA

FIGURE 41

MGSVLGLCSMASWIPCLCGSAPCLLCRCCPSGNNSTVTRLIYALFLLVGVCVACVMLIPGME
EQLNKIPGFCENEKGVVPCNILVGYKAVYRLCFGLAMFYLLLSLLMIKVKSSSDPRAAVHNG
FWFFKFAAAIAIIIGAFFIPEGTFTTVWFYVGMAGAFCFILIQLVLLIDFAHSWNESWVEKM
EEGNSRCWYAALLSATALNYLLSLVAIVLFFVYYTHPASCSENKAFISVNMLLCVGASVMSI
LPKIQESQPRSGLLQSSVITVYTMylTWSAMTNEPETNCNPSLLSIIGYNTTSTVPKEGQSV
QWWHAQGIIGLILFLLCVFYSSIRTSNNSQVNKLTLTSDESTLIEDGGARSDGSLEDGDDVH
RAVDNERDGVITYSYSFFHFMLFLASLYIMMTLTNWSRYEPSREMKSQWTAVVVKISSSWIGI
VLYVWTLVAPLVLTNRDFD

090922 220900

CGCAGAAAGAAGCTGTCTCCATCTTGTCTGTATCCCGCTGCTTCTTGNGACGTTGTGGAGAT
GGGGAGCGTCCCTGGGGCTGTGCTCCATGGCGAGCTGGATACCATGTTTGTGTGGAAGTGCC
CCGTGTTTGCTATGCCGATGCTGTCCTAGTGGAACAANTCCACTGTAAC TAGATTGATCTA
TGCACTTTTCTTGCTTGTTGGAGTATGTGTAGCTTGTGTAATGTTGATACCAGGAATGGAAG
AACAACTGAATAAGATTCCTGGATTTTGTGAGAATGAGAAAGGTGTTGTCCCTTGTAACATT
TTGGTTGGCTATAAAGCTGTATATCGTTTGTGCTTTGGTTTGGCTATGTTCTATCTTCTTCT
CTCTTTACTAATGATCAAAGTGAAGAGTAGCAGTGATCCTAGAGCTGCAGTGCACAATGGAT
TTTGGTTCTTTAAATTTGCTGCAGCAATTGCAATTATTATTGGGGC

FIGURE 43

GTTATTGTGAACTTTGTGGAGATGGGAGGTCNTGGGGCTGTGTTCCATGGCGAGCTGGATAC
CANGTTTGTGTGGAAGTGCCCCGTGTTTGNATGCCGATGCTGTCCTAGTGGAACAANTCC
ACTGTAATTAGATTGATNTATGCACTTTTNTTGCTTGTTGGAGTANGTGTAGCTTGTGTAAT
GTTGATACCAGGAATGGAAGAACAACCTGAATAAGATTCCTGGATTTTGTGAGAATGAGAAAG
GTGTTGTCCCTTGTAACATTTTGGTTGGCTATAAAGCTGTATATNGTTTGTGCTTTGGTTTG
GCTANGTTCTATNTTCTTCTCTCTTTACTAATGATCAAAGTGAAGAGTAGCAGTGATCCTAG
AGCTGCAGTGCACAATGGATTTTGGTTTTTTAAATTTGCTGCAGCAATTGCAATTATTATTG
GGGC

0997226560

FIGURE 44

AAGAAGCTGTCTCCATCTTGTCTGTATCCGCTGCTCTTGTGAACGTTNTGGAGATGGGGAGC
GTCCTTGGGGTTGTGCTCCATGGCGAGCTGGATACCATGTTTGTGTGGAAGTGCCCCGTGTT
TGCTATGCCGATGCTGTCCTAGTGGAACAACCTCCACTGTAAC TAGATTGATCTATGCACTT
TTCTTGCTTGTTGGAGTATGTGTAGCTTGTGTAATGTTGATACCAGGAATGGAAGAACAAC
GAATAAGATTCCCTGGATTTTGTGAGAAATGAGAAAGGTGTTGTCCCTTGTAACATTTTGGTTG
GCTATAAAGCTGTATATCGTTTGTGCTTTGGTTTGGCTATGTTCTATCTTCTTCTCTTTA
CTAATGATCAAAGTGAAGAGTAGCAGTGATCCTAGAGCTGCAGTGCACAATGGATTTTGGTT
CTTTAAATTTGCTGCAGCAATTGCAATTATTATTGGGGC

TGTCTTGGGGTTGTGCTCCATGGCGAGCTGGATACCATGTTTGTGTGGAAGTGCCCCGTGTT

GCTGTCCTTAGTGGAAACAANTCCAACCTTGTAACCTTGGATTGATCTATGCACTTTTTCCCTTG
CTTGTTGGAGTATGTGTAGCTTTGTGTAATGTTGTTCCCAGGATTGGANGAACAACTGAATA
AGATTCCCTGGATTTTTGTGAGAATGAGAAAGGTGTTGTCCCCTTGTAACATTTTTGGTTGGC
TATAAAGCTGTATATCGTTTGTGCTTTGGTTTGGCTATGTTCTATCTTCTTCTCTTTACT
AATGATCAAAGTGAAGAGTAGCAGTGATCCTAGAGCTGCAGTGCACAATGGATTTTGGTTCT
TTAAATTTGCTGCAGCAATTGCAATTATTATTGGGGCATTCTTCATTCCAGAAGGAACTTTT
ACAACTGTGTGGTTTTATGTAGGCATGGCAGGTGCCTTTTGTTCATCCTCATACAAGTAGT
CTTACTTATTGATTTTGCACATTCATGGAATGAATCGTGGGTTGAAAAAATGGAAGAAGGGA
ACTCGAGATGTTGGTATGCAGCCTTGTTATCAGCTACAGCTCTGAATTATCTGCTGTCTTTA
GTTGCTATCGTCCTGTTCTTTGTCTACTACACTCATCCAGCCAGTTGTTTCAGAAAACAAGGC
GTTTCATCAGTGTCAACATGCTCCTCTGCGTTGGTGCTTCTGTAATG

FIGURE 46

CTCGGGCGCGCACAGGCAGCTCGGTTTGGCCCTGCGATTGAGCTGCGGGTTCGCGGCCGGCGCCGGCCCTCTCCAAT
GGCAAATGTGTGTGGCTGGAGGCGAGCGCGAGGCTTTCGGCAAAGGCAGTCGAGTGTTTGCAGACCGGGGCGAG
TCCTGTGAAAGCAGATAAAAGAAAACATTTATTAACGTGTCATTACGAGGGGAGCGCCCGGCCGGGGCTGTGCGC
ACTCCCCCGGGAACATTTGGCTCCCTCCAGCTCCGAGAGAGGAGAAGAAGAAAGCGGAAAAGAGGCAGATTAC
GTCGTTTTCCAGCCAAAGTGGACCTGATCGATGGCCCTCCTGAATTTATCACGATATTTGATTTATTAGCGATGCC
CCCTGGTTTGTGTGTACGCACACACACGTGCACACAAGGCTCTGGCTCGCTTCCCTCCCTCGTTTTCCAGCTCC
TGGGCGAATCCACATCTGTTTCAACTCTCCGCCGAGGGCGAGCAGGAGCGAGAGTGTGTGCAATCTGCGAGTG
AAGAGGGACGAGGGAAAAGAAACAAAGCCACAGACGCAACTTGAGACTCCCGCATCCCAAAAGAAGCACCAGAT
CAGCAAAAAAAGAAGATCGGGCCCCCGAGCCTCGTGCTGTGCTTGTGCTCCGCAACTGTGTTCTCCCTGCTGGG
TGGAAGCTCGGCCCTTCTGTGCGACCACCGCCTGAAAGGCAGGTTTTCAGAGGGACCGCAGGAACATCCGCCCA
ACATCATCCTGGTGCTGACGGACGACAGGATGTGGAGCTGGGTTCCATGCAGGTGATGAACAAGACCCGGCGC
ATCATGGAGCAGGGCGGGGCGCACTTCATCAACGCCCTTCGTGACCACACCCATGTGCTGCCCTCAGCCTCCTC
CATCCTCACTGGCAAGTACGTCCACAACCACAACACCTACACCAACAATGAGAACTGCTCCTCGCCCTCCTGGC
AGGCACAGCAGCAGAGCCGACCTTTGCCGTGTACCTCAATAGCACTGGCTACCGGACAGCTTTCTTCGGGAG
TATCTTAATGAATACAACGGCTCCTACGTGCCACCCGGCTGGAAGGAGTGGGTTCGGACTCCTTAAAACTCCCC
CTTTTATAACTACAGCTGTGTGGAACGGGGTGAAGAGAAGCACGGCTCCGACTACTCCAAGGATTACCTCA
CAGACCTCATCACCAATGACAGCGTGAGCTTCTTCGCACTTCCAAGAAGATGTACCCGCACAGGCCAGTCCCTC
ATGGTCATCAGCCATGCAGCCCCCACGGCCCTGAGGATTACGCCCCACAATATTACGCCTCTTCCCAACGC
ATCTCAGCACATCAGCCGAGCTACAACCTACGCGCCCAACCCGGACAAACACTGGATCATGCGCTACACGGGGC
CCATGAAGCCCATCCACATGGAATTCACCAACATGCTCCAGCGGAAGCGCTTGCAGACCCTCATGTGCTGGAC
GACTCCATGGAGACGATTTACAACATGCTGGTTGAGACGGGCGAGCTGGACAACACGTACATCGTATACACCGC
CGACCACGGTTACCACATCGGCCAGTTTGGCCTGGTGAAAGGGAAATCCATGCCATATGAGTTTGACATCAGGG
TCCCGTTCTACGTGAGGGGCCCCAACGTGGAAGCCGGCTGTCTGAATCCCCACATCGTCTCAACATTGACCTG
GCCCCACCATCCTGGACATTGCAGGCCTGGACATACCTGCGGATATGGACGGGAAATCCATCCTCAAGCTGCT
GGACACGGAGCGGCCGCTGAATCGGTTTCACTTGAAAAAGAAGATGAGGGTCTGGCGGGACTCCTTCTTGGTGG
AGAGAGGCAAGCTGCTACACAAGAGAGACAATGACAAGGTGGACGCCAGGAGGAGAATTTCTGCCCAAGTAC
CAGCGTGTGAAGGACCTGTGTGAGCGTGTGAGTACCAGACGGCGTGTGAGCAGCTGGGACAGAAGTGGCAGTG
TGTGGAGGACGCCACGGGGAAGCTGAAGCTGCATAAGTGCAAGGGCCCCATGCGGCTGGGCGGCAGCAGAGCCC
TCTCCAACCTCGTGCCCAAGTACTACGGGCAGGGCAGCGAGGCCTGCACCTGTGACAGCGGGGACTACAAGCTC
AGCCTGGCCGACGCCGGAATAACTCTTCAAGAAGAAGTACAAGGCCAGCTATGTCCGCAGTCGCTCCATCCG
CTCAGTGGCCATCGAGGTGGACGGCAGGGTGTACCACGTAGGCCTGGGTGATGCCGCCAGCCCCGAAACCTCA
CCAAGCGCACTTGCCAGGGGCCCCCTGAGGACCAAGATGACAAGGATGGTGGGACTTCAGTGGCACTGGAGGC
CTTCCCAGCTACTCGCCGCAACCCCATTAAGGTGACACATCGTGCTACATCCTAGAGAACGACACAGTCCA
GTGTGACCTGGACCTGTACAAGTCCCTGCAGGCCTGGAAGACCACAAGCTGCACATCGACCACGAGATTGAAA
CCCTGCAGAACAAAATTAAGAACCTGAGGGAAGTCCGAGGTACCTGAAGAAAAAGCGGCCAGAAGATGTGAC
TGTCACAAAATCAGCTACCACACCCAGCACAAAGGCCGCTCAAGCACAGAGGCTCCAGTCTGCATCCTTTAG
GAAGGGCCTGCAAGAGAAGGACAAGGTGTGGCTGTTGCGGGAGCAGAAGCGCAAGAAGAACTCCGCAAGCTGC
TCAAGCGCCTGCAGAACACGACACGTGCAGCATGCCAGGCCTCACGTGCTTCACCCACGACAACGACACTGG
CAGACGGCGCCTTCTGGACACTGGGGCCTTCTGTGCTGCACCGCCCAACAATAACACGTACTGGTGCAT
GAGGACCATCAATGAGACTCACAATTTCTCTTCTGTGAATTTGCAACTGGCTTCTTAGAGTACTTTGATCTCA
ACACAGACCCCTACCAGCTGATGAATGCAGTGAACACACTGGACAGGGATGTCTCAACCAGCTACACGTACAG
CTCATGGAGCTGAGGAGCTGCAAGGGTTACAAGCAGTGTAAACCCCGGACTCGAAACATGGACCTGGATGGAGG
AAGCTATGAGCAATACAGGCAGTTTCAGCGTCGAAAGTGGCCAGAAATGAAGAGACCTTCTTCCAAATCACTGG
GACAACTGTGGGAAGGCTGGGAAGGTTAAAGAAAACAACAGAGGTGGACCTCCAAAAACATAGAGGCATCACCTGA
CTGCACAGGCAATGAAAAACCATGTGGGTGATTTCCAGCAGACCTGTGCTATTGGCCAGGAGGCTGAGAAAAGC
AAGCACGCACTCTCAGTCAACATGACAGATTCTGGAGGATAACCAGCAGGAGCAGAGATAACTTCAGGAAGTCC
ATTTTTGCCCCGCTTTTGGCTTTGGATTATACCTCACCAGCTGCACAAAATGCATTTTTTCGTATCAAAAAGTC
ACCATAACCTCCCCCAGAAGCTCACAAAGGAAAACGGAGAGAGCGAGAGAGAGATTTCTTGGAAATTTCT
TCCCAAGGGCGAAAGTCATTGGAATTTTAAATCATAGGGGAAAAGCAGTCTCTTAAATCCTCTTATTCTT
TTGGTTTGTCAAAAGGAACATAAGAGCAGGACAGAGCAACGTGGAGAGGCTGAAAACAGTGCAGAGACG
TTTGACAATGAGTCAGTAGCACAAAAGAGATGACATTTACCTAGCACTATAAACCTGGTTGCTCTGAAGAAA
CTGCCTTCAATTGTATATATGTGACTATTTACATGTAATCAACATGGGAACTTTTAGGGGAACCTAATAAGAAAT
CCCAATTTTCAGGAGTGGTGGTGTCAATAAACGCTCTGTGGCCAGTGTAAAAGAAAAA

0989727-11901

FIGURE 47

MGPPSLVLCCLLSATVFSLLGGSSAFLSHHRLKGRFQDRRNIRPNIILVLTDDQDVELGSMQ
VMNKTRRIMEQGGAHFINAFVTTMCCPSRSSILTGKYVHNHNTYTNNENCSSPSWQAQHE
RTFAVYLNSTGYRTAFFGKYLNEYNGSYVPPGWKEWVGLLKNSRFYNYTLCRNGVKEKHGSD
YSKDYLTDLITNDSVSFFRTSKKMYPHRPVLMVISHAAPHGPEDSAPQYSRLFPNASQHITP
SYNYAPNPDKHWIMRYTGPMKPIHMEFTNMLQRKRLQTLMSVDDSMETIYNMLVETGELDNT
YIVYTADHGYHIGQFGLVKGKSMPYEFDIRVPFYVRGPNVEAGCLNPHIVLNIDLAPTILDI
AGLDIPADMKGKLSILKLLDTERPVRNRFHLKKKMRVWRDSFLVERGKLLHHRDNDKVDAQEEN
FLPKYQRVKDLQRAEYQTACEQLGQKWQCVEDATGKLKLHKCKGPMRLGGSRALSNLVPKY
YGQGSEACTCDSGDYKLSLAGRRKKLFKKKKYKASYVRSRSIRSV AIEVDGRVYHVGLGDAAQ
PRNLTKRHWPGAPEDQDDKDGGDFSGTGGLPDYSAANPIKVTHRCYILENDTVQCDLDLYKS
LQAWKDHKLHIDHEIETLQNKIKNLREVRGHLKKKRPEECDCHKISYHTQHKGRLKHRGSSL
HPFRKGLQEKDKVWLLREQKRKKLRKLLKRLQNNDTCSMPGLTCFTHDNQHWQTAPFWTLG
PFCACTSANNNTYWCMRTINETHNLFCEFATGFLEYFDLNTDPYQLMNAVNTLDRDVLNQL
HVQLMELRSCKGYKQCNPRTNMDLDGGSYEQYRQFQRRKWPEMKRPSSKSLGQLWEGWEG

FIGURE 48

AACAAAGTTCAGTGACTGAGAGGGCTGAGCGGAGGCTGCTGAAGGGGAGAAAGGAGTGAGGA
GCTGCTGGGCAGAGAGGGACTGTCCGGCTCCCAGATGCTGGGCCTCCTGGGGAGCACAGCCC
TCGTGGGATGGATCACAGGTGCTGCTGTGGCGGTCCTGCTGCTGCTGCTGCTGCTGGCCACC
TGCCTTTTCCACGGACGGCAGGACTGTGACGTGGAGAGGAACCGTACAGCTGCAGGGGGAAA
CCGAGTCCGCCGGGGCCAGCCTTGGCCCTTCCGGCGGCGGGGCCACCTGGGAATCTTTCACC
ATCACCGTCATCCTGGCCACGTATCTCATGTGCCGAATGTGGGCCTCCACCACCACCACCAC
CCCCGCCACACCCCTCACCACCTCCACCACCACCACCACCCCCACCGCCACCATCCCCGCCA
CGCTCGCTGAGGCTGCTGTGCGCCGGTGCCTGTGGACAGCAGCTGCCCCTGCCCTCCCATCTG
TTCCCAGGACAAGTGGACCCCATGTTTCCATGTGGAAGGATGCATCTCTGGGGTGAACGAGG
GGAACAATAGACTGGGGCTTGCTCCAGCTGCATTTGCATGGCATGCCCCAGTGTACTATGGC
AGCAGAGAATGGAGGAACACTGGGTCTGCAGTGCTGAAGGGTTTGGGGAGTGGAGAGCAAGG
GTGCTCTTTCGGGGCTGGACAGCCCGTCTTGTGACAGTGAAGTCCCAGTGAGCCCCAGAAATG
ACAAGCGTGTCTTGGCAGAGCCAGCACACAAGTGGATGTGAAGTGCCCGTCTTGACCTCCTC
ATCAGGCTGCTGCAGGCCTCTGGCGGGCAGGGCACTGGGAGAGGCCCTGAGAATGTCCTTTT
GGTTTGGAGAAGGCAGTGTGAGGCTGCACAGTCAATTCATCGGTGCCTTAGTCCAAGAAAAT
AAAAACCACTAAGAAGCTTTAAAAAAAAAAAAAAAAAAAAA

FIGURE 49

MLGLLGSTALVGWITGAAVAVLLLLLLLLLATCLFHGRQDCDVERNRTAAGGNRVVRAQPWPER
RRGHLGIFHHHRHPGHVSHVPNVGLHHHHHPRHTPHHLHHHHHPRHHPRHAR

FIGURE 50

GGCGGCTGCTGAGCTGCCTTGAGGTGCAGTGTTGGGGATCCAGAGCCATGTCGGACCTGCTA
CTACTGGGCCTGATTGGGGGCCTGACTCTCTTACTGCTGCTGACGCTGCTGGCCTTTGCCGG
GTACTCAGGGCTACTGGCTGGGGTGGAAGTGAGTGCTGGGTCACCCCCCATCCGCAACGTCA
CTGTGGCCTACAAGTTCCACATGGGGCTCTATGGTGAGACTGGGCGGCTTTTCACTGAGAGC
TGCAGCATCTCTCCCAAGCTCCGCTCCATCGCTGTCTACTATGACAACCCCCACATGGTGCC
CCCTGATAAGTGCCGATGTGCCGTGGGCAGCATCCTGAGTGAAGGTGAGGAATCGCCCTCCC
CTGAGCTCATCGACCTCTACCAGAAATTTGGCTTCAAGGTGTTCTCCTTCCCGGCACCCAGC
CATGTGGTGACAGCCACCTTCCCCTACACCACCATTTCTGTCCATCTGGCTGGCTACCCGCCG
TGTCCATCCTGCCTTGACACCTACATCAAGGAGCGGAAGCTGTGTGCCTATCCTCGGCTGG
AGATCTACCAGGAAGACCAGATCCATTTTATGTGCCCACTGGCACGGCAGGGAGACTTCTAT
GTGCCTGAGATGAAGGAGACAGAGTGGAATGGCGGGGGGCTTGTGGAGGCCATTGACACCCA
GGTGGATGGCACAGGAGCTGACACAATGAGTGACACGAGTTCTGTAAGCTTGGAAGTGAGCC
CTGGCAGCCGGGAGACTTCAGCTGCCACACTGTCACCTGGGGCGAGCAGCCGTGGCTGGGAT
GACGGTGACACCCGCAGCGAGCACAGCTACAGCGAGTCAGGTGCCAGCGGCTCCTCTTTTGA
GGAGCTGGACTTGAGGGGCGAGGGGCCCTTAGGGGAGTCACGGCTGGACCCTGGGACTGAGC
CCCTGGGGACTACCAAGTGGCTCTGGGAGCCCACTGCCCCTGAGAAGGGCAAGGAGTAACCC
ATGGCCTGCACCCTCCTGCAGTGCAGTTGCTGAGGAACTGAGCAGACTCTCCAGCAGACTCT
CCAGCCCTCTTCCTCCTTCTCTGGGGGAGGAGGGGTTCTGAGGGACCTGACTTCCCCTGC
TCCAGGCCTCTTGCTAAGCCTTCTCCTCACTGCCCTTTAGGCTCCCAGGGCCAGAGGAGCCA
GGGACTATTTTCTGCACCAGCCCCCAGGGCTGCCGCCCCTGTTGTGTCTTTTTTTCAGACTC
ACAGTGGAGCTTCCAGGACCCAGAATAAAGCCAATGATTTACTTGTTTCACCTGGAAAAAA
AAAAA

MSDLLLLGLIGGLTLLLLLTLLAFAGYSGLLAGVEVSAGSPPIRNVTVAYKFHMGLYGETGR
LFTESCSISPKLRSIAVYYDNPHMVPPDKRCRAVGSILSEGEESPSPELIDLYQKFGFKVFS
FPAPSHVVTATFPYTTILSIWLATRRVHPALDTYIKERKLCAYPRLEIYQEDQIHFMCPLAR
QGDFYVPEMKETEWKWRGLVEAIDTQVDGTGADTMSDTSSVSLEVSPGSRETSAAATLSPGAS
SRGWDDGDTRSEHSYSESGASGSSFEELDLEGEGLGESRLDPGTEPLGTTKWLWEPTAPEK
GKE

[illegible]

CCGCGGGAACGCTGTCTGGCTGCCGCCACCCGAACAGCCTGTCTCTGGTGCCCCGGCTCCCT
GCCCCGCGCCAGTCAATGACCCTGCGCCCCCTCACTCCTCCCGCTCCATCTGCTGCTGCTGCT
GCTGCTCAGTGCGGCGGTGTGCCGGGCTGAGGCTGGGCTCGAAACCGAAAGTCCCGTCCGGA
CCCTCCAAGTGAGACCCTGGTGGAGCCCCAGAACCATGTGCCGAGCCCGCTGCTTTTGA
GACACGCTTCACATACTACACGGGAAGCTTGGTAGATGGACGTATTATTGACACCTCCCT
GACCAGAGACCCTCTGGTTATAGAACTTGGCCAAAAGCAGGTGATTCCAGGTCTGGAGCAGA
GTCTTCTCGACATGTGTGTGGGAGAGAAGCGAAGGGCAATCATTCTTCTCACTTGGCCTAT
GGAAAACGGGGATTTCCACCATCTGTCCCAGCGGATGCAGTGGTGCAGTATGACGTGGAGCT
GATTGCACTAATCCGAGCCAACTACTGGCTAAAGCTGGTGAAGGGCATTTTGCCTCTGGTAG
GGATGGCCATGGTGCCAGCCCTCCTGGGCCTCATTTGGGTATCACCTATACAGAAAGGCCAAT
AGACCCAAAGTCTCCAAAAGAAGCTCAAGGAAGAGAAACGAAACAAGAGCAAAAAGAAATA
ATAAATAATAAATTTTAAAAAACTTAAAAAAAAAAAAAAAAAAAA

FIGURE 53

MTLRPSLLPLHLLLLLLLLSAAVCRAEAGLETESPVRTLQVETLVEPPEPCAEPAAFCDTLHI
HYTGSLVDGRIIDTSLTRDPLVIELGQKQVIPGLEQSLDMCVGEKRRAIIPSHLAYGKRGF
PPSVPADAVVQYDVELIALIRANYWLKLVKGILPLVGMAMVPALLGLIGYHLYRKANRPKVS
KKKLKEEKRNKSKKK

bioRxiv preprint doi: <https://doi.org/10.1101/2020.03.26.000000>; this version posted March 26, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 55

CCGAAAGTCCCGTCCGGACCCTCCAAGTGGAGACCCTGGTGGAGCCCCCAGAACCATGTGCC
GAGCCCGCTGCTTTTGGAGACACGCTTCACATACACTACACGGGAAGCTTGGTAGATGGACG
TATTATTGACACCTCCCTGACCAGAGACCCTCTGGTTATAGAACTTGGCCAAAAGCAGGTGA
TTCCAGGTCTGGAGCAGAGTCTTCTCGACATGTGTGTGGGAGAGAAGCGAAGGGCAATCATT
CCTTCTCACTTGGCCTATGGAAAACGGGGATTTCCACCATCTGTCCCAGCGGATGCAGTGGT
GCAGTATGACGTGGAGCTGATTGCACTAATCCGAGCCAACTACTGGCTAAAGCTGGTGAAGG
GCATTTTGCCTCTGGTAGGGATGGCCATGGTGCCAGCCCTCCTGGGCCTCATTGGGTATCAC
CTATACAGAAAGGCCAATAGACCCAAAGTCTCCAAAAAGAAGCTCAAGGAAGAGAAACGAAA
CAAGAGCAAAAAGAAATAATAAATAATAAATTTTAAAAAACTTAAAA

57/330

FIGURE 57

MSGFLEGLRCSECIDWGEKRNTIASIAAGVLFFTGWIIIDA AVIYPTMKDFNHSYHACGVI
ATIAFLMINAVSNGQVRGDSYSEGCLGQTGARIWLFVGFMLAFGSLIASMWILFGGYVAKEK
DIVYPGIAVFFQNAFIFFGGLVFKFGRTEDLWQ

bioRxiv preprint doi: <https://doi.org/10.1101/2020.03.10.332660>; this version posted March 10, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 59

TGGACGGACCTGAAAAAATGTTTGGATTNTAGAGGGNTTGAGATGTTCAGAATGCATGAC
TGGGGGAAAAGCGCAAATACTATTGCTTCCATTGCTGCTGGTGTANTATTTTTTACAGGCTG
GTGGATTATCATAGATGCAGNTGTTATTTATCCCACCATGAAAGATTTCAACCANTCATACC
ATGCCTGTGGTGTATAGCAACCATAGCCTTCNTAATGATTAATGCAGTATCGAATGGACAA
GTCCGAGGTGATAGTTACAGTGAAGGTTGTTTGGGTCAAACAGGTGCTCGCATTTGGCTTTT
CGTTGGTTTCATGTTGGCCTTTGGATCTCTGATTGCATCTATGTGGATTCTTTTTGGAGGTT
ATGTTGCTAAAGAAAAAGACATAGTATACCCTGGAATTGNTGTATTTTTCCAGAATGCCTTC
ATCTTTTTTTGGAGGGCTGGTTTTTAAGTTTGGCCGCACTGAAGANTTATGGCAGTG

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 61

TAGAGGGCTTGAGATGCTCAGAATGCATTGACTGGGGGGAAAAGCGCAATANTATTGCTTCC
ATTGNTGNTGGTGTANTATTTTTTTTACAGGCTGGTGGATTATNATAGATGCAGCTGTTATTT
ATCCCACCATGAAAGATTTNAACCANTCATACCATGCCTGTGGTGTTATAGCAACCATAGCC
TTCCTAATGATTAATGCAGTATNGAATGGACAAGTCCGAGGTGATAGTTACAGTGAAGGTTG
TTTGGGTCAAACAGGTGNTNGCATTGCGCTTTTNGTTGGTTTCATGTTGGCCTTTGGATCTN
TGATTGCATTTATGTGGATTNTTTTTTGGAGGTTATGTTGCTAAAGNAAAAGACATAGTATAC
CCTGT

FIGURE 63

CGACGCCGGCGTG**ATG**TGGCTTCCGCTGGTGCTGCTCCTGGCTGTGCTGCTGCTGGCCGTCC
TCTGCAAAGTTTACTTTGGGACTATTCTCTGGCAGCTCCCCGAATCCTTTCTCCGAAGATGTC
AAACGGCCCCCAGCGCCCCCTGGTAACTGACAAGGAGGCCAGGAAGAAGGTTCTCAAACAAGC
TTTTTTCAGCCAACCAAGTGCCGGAGAAGCTGGATGTGGTGGTAATTGGCAGTGGCTTTGGGG
GCCTGGCTGCAGCTGCAATTCTAGCTAAAGCTGGCAAGCGAGTCTGGTGCTGGAACAACAT
ACCAAGGCAGGGGGGCTGCTGTCATACCTTTGGAAAGAATGGCCTTGAATTTGACACAGGAAT
CCATTACATTGGGCGTATGGAAGAGGGCAGCATTGGCCGTTTTATCTTGGACCAGATCACTG
AAGGGCAGCTGGACTGGGCTCCCCTGTCCTCTCCTTTTGGACATCATGGTACTGGAAGGGCCC
AATGGCCGAAAGGAGTACCCCATGTACAGTGGAGAGAAAGCCTACATTCAGGGCCTCAAGGA
GAAGTTTCCACAGGAGGAAGCTATCATTGACAAGTATATAAAGCTGGTTAAGGTGGTATCCA
GTGGAGCCCCCTCATGCCATCCTGTTGAAATTCCTCCCATTTGCCGTGGTTTCACTCCTCGAC
AGGTGTGGGCTGCTGACTCGTTTCTCTCCATTCTTCAAGCATCCACCCAGAGCCTGGCTGA
GGTCTGCAGCAGCTGGGGGCTCCTCTGAGCTCCAGGCAGTACTCAGCTACATCTTCCCCA
CTTACGGTGTCAACCCCAACCCACAGTGCCTTTTCCATGCACGCCCTGCTGGTCAACCACTAC
ATGAAAGGAGGCTTTTTATCCCCGAGGGGGTTCAGTGAAATTGCCTTCCACACCATCCCTGT
GATTCAGCGGGCTGGGGGCGCTGTCCTCACAAAGGCCACTGTGCAGAGTGTGTTGCTGGACT
CAGCTGGGAAAGCCTGTGGTGTGAGTGTGAAGAAGGGGCATGAGCTGGTGAACATCTATTGC
CCCATCGTGGTCTCCAACGCAGGACTGTTCAACACCTATGAACACCTACTGCCGGGGAACGC
CCGCTGCCTGCCAGGTGTGAAGCAGCAACTGGGGACGGTGCGGCCCGGCTTAGGCATGACCT
CTGTTTTTCTGCTGCGAGGCACCAAGGAAGACCTGCATCTGCCGTCCACCAACTACTAT
GTTTACTATGACACGGACATGGACCAGGCGATGGAGCGCTACGTCTCCATGCCAGGGAAGA
GGCTGCGGAACACATCCCTCTTCTCTTCTTCCATCAGCCAAAGATCCGACCTGGG
AGGACCGATTCCCAGGCCGGTCCACCATGATCATGCTCATACCCACTGCCTACGAGTGGTTT
GAGGAGTGGCAGGCGGAGCTGAAGGGAAGCGGGGCAGTGACTATGAGACCTTCAAAAACCTC
CTTTGTGGAAGCCTCTATGTGAGTGGTCTGAAACTGTTCCACAGCTGGAGGGGAAGGTGG
AGAGTGTGACTGCAGGATCCCCACTCACCACACAGTCTATCTGGCTGCTCCCCGAGGTGCC
TGCTACGGGGCTGACCATGACCTGGGCCGCTGCACCCTTGTGTGATGGCCTCCTTGAGGGC
CCAGAGCCCCATCCCCAACCTCTATCTGACAGGCCAGGATATCTTACCTGTGGACTGGTGC
GGGCCCTGCAAGGTGCCCTGCTGTGCAGCAGCGCCATCCTGAAGCGGAACCTTGACTCAGAC
CTTAAGAATCTTGATTCTAGGATCCGGGCACAGAAGAAAAAGAAT**TAG**TTCCATCAGGGAGG
AGTCAGAGGAATTTGCCCAATGGCTGGGGCATCTCCCTTGACTTACCCATAATGTCTTTCTG
CATTAGTTCCTTGACGTATAAAGCACTCTAATTTGGTTCTGATGCCTGAAGAGAGGCCTAG
TTTAAATCACAATTCGAATCTGGGGCAATGGAATCACTGCTTCCAGCTGGGGCAGGTGAGA
TCTTTACGCCTTTTATAACATGCCATCCCTACTAATAGGATATTGACTTGGATAGCTTGATG
TCTCATGACGAGCGGCGCTCTGCATCCCTCACCCATGCCTCCTAACTCAGTGATCAAAGCGA
ATATTCCATCTGTGGATAGAACCCCTGGCAGTGTGTCAGCTCAACCTGGTGGGTTTCAAGTTC
TGTCCTGAGGCTTCTGCTCTCATTCAATTTAGTGCTACGCTGCACAGTTCTACACTGTCAAGG
GAAAAGGGAGACTAATGAGGCTTAACCAAAACCTGGGCGTGGTTTTGGTTGCCATTCCATA
GGTTTGGAGAGCTCTAGATCTCTTTTGTGCTGGGTTTCAAGTGGCTCTTCAAGGGGACAGGAAAT
GCCTGTGTCTGGCCAGTGTGGTTCTGGAGCTTTGGGGTAACAGCAGGATCCATCAGTTAGTA
GGGTGCATGTGAGATGATCATATCCAATTCATATGGAAGTCCCGGGTCTGTCTTCTTATCA
TCGGGGTGGCAGCTGGTTCTCAATGTGCCAGCAGGGACTCAGTACCTGAGCCTCAATCAAGC
CTTATCCACCAATAACACAGGGAAGGGTGATGCAGGGAAGGGTGACATCAGGAGTCAGGGCA
TGGACTGGTAAGATGAATACTTTGCTGGGCTGAAGCAGGCTGCAGGGCATTCAGCCAAGGG
CACAGCAGGGGACAGTGCAGGGAGGTGTGGGGTAAGGGAGGGAAGTCACATCAGAAAAGGGA
AAGCCACGGAATGTGTGTGAAGCCCAGAAATGGCATTGTGAGTTAATTAGCACATGTGAGGG
TTAGACAGGTAGGTGAATGCAAGCTCAAGGTTTGGAAAAATGACTTTTCAGTTATGTCTTTG
GTATCAGACATACGAAAGTCTCTTTGTAGTTCTGTGTTAATGTAACATTAATAAATTTATTG
ATTCCATTGCTTTAAAAA

FIGURE 65

GCAGCGGCGAGGCGGCGGTGGTGGCTGAGTCCGTGGTGGCAGAGGCGAAGGCGACAGCTCTA
GGGGTTGGCACCGGCCCGAGAGGAGG**ATG**CGGGTCCGGATAGGGCTGACGCTGCTGCTGTG
TGCGGTGCTGCTGAGCTTGGCCTCGGCGTCCTCGGATGAAGAAGGCAGCCAGGATGAATCCT
TAGATTCCAAGACTACTTTGACATCAGATGAGTCAGTAAAGGACCATACTACTGCAGGCAGA
GTAGTTGCTGGTCAAATATTTCTTGATTGAGAAGAATCTGAATTAGAATCCTCTATTCAAGA
AGAGGAAGACAGCCTCAAGAGCCAAGAGGGGGAAAGTGTCACAGAAGATATCAGCTTTCTAG
AGTCTCCAAATCCAGAAAACAAGGACTATGAAGAGCCAAAGAAAGTACGGAAACCAGCTTTG
ACCGCCATTGAAGGCACAGCACATGGGGAGCCCTGCCACTTCCCTTTTCTTTTCTTAGATAA
GGAGTATGATGAATGTACATCAGATGGGAGGGAAGATGGCAGACTGTGGTGTGCTACAACCT
ATGACTACAAAGCAGATGAAAAGTGGGGCTTTTGTGAACTGAAGAAGAGGCTGCTAAGAGA
CGGCAGATGCAGGAAGCAGAAATGATGTATCAAATGGAATGAAAATCCTTAATGGAAGCAA
TAAGAAAAGCCAAAAAAGAGAAGCATATCGGTATCTCCAAAAGGCAGCAAGCATGAACCATA
CCAAAGCCCTGGAGAGAGTGTGATATGCTCTTTTATTTGGTGATTACTTGCCACAGAATATC
CAGGCAGCGAGAGAGATGTTTGAGAAGCTGACTGAGGAAGGCTCTCCCAAGGGACAGACTGC
TCTTGGCTTTTCTGTATGCCTCTGGACTTGGTGTAAATTCAGTCAGGCAAAGGCTCTTGAT
ATTATACATTTGGAGCTCTTGGGGCAATCTAATAGCCACATGGTTTTGGTAAGTAGACTT
TAGTGGAAGGCTAATAATATTAACATCAGAAGAATTTGTGGTTTATAGCGGCCACAACTTT
TCAGCTTTCATGATCCAGATTTGCTTGTATTAAGACCAAATATTCAGTTGAACTTCCTTCAA
ATTCTTGTTAATGGATATAACACATGGAATCTACATGTAAATGAAAGTTGGTGGAGTCCACA
ATTTTTCTTTAAATGATTAGTTTGGCTGATTGCCCTAAAAAGAGAGATCTGATAAATGGC
TCTTTTTAAATTTTCTCTGAGTTGGAATTGTCAGAATCATTTTTTACATTAGATTATCATAA
TTTTAAAAATTTTCTTTAGTTTTTCAAATTTTGTAAATGGTGGCTATAGAAAAACAACAT
GAAATATTATACAATATTTTGCAACAATGCCCTAAGAATTGTTAAATTCATGGAGTTATTT
GTGCAGAATGACTCCAGAGAGCTCTACTTTCTGTTTTTTACTTTTCATGATTGGCTGTCTTC
CCATTTATTCTGGTCATTTATTGCTAGTGACACTGTGCCTGCTTCCAGTAGTCTCATTTTCC
CTATTTTGCTAATTTGTTACTTTTTCTTTGCTAATTTGGAAGATTAACATTTTTTAATAAA
ATTATGTCTAAGATTAAAAA
AA
AA

FIGURE 67

CTTCCCAGCCCTGTGCCCCAAAGCACCTGGAGCATATAGCCTTGCAGAACTTCTACTTGCCT
GCCTCCCTGCCTCTGGCC**ATG**GCCTGCCGGTGCCTCAGCTTCCTTCTGATGGGGACCTTCCT
GTCAGTTTCCCAGACAGTCCTGGCCCAGCTGGATGCACTGCTGGTCTTCCCAGGCCAAGTGG
CTCAACTCTCCTGCACGCTCAGCCCCAGCACGTACCATCAGGGACTACGGTGTGTCCTGG
TACCAGCAGCGGGCAGGCAGTGGCCCTCGATATCTCCTCTACTACCGCTCGGAGGAGGATCA
CCACCGGCCTGCTGACATCCCCGATCGATTCTCGGCAGCCAAGGATGAGGCCCAATGCCT
GTGTCCTCACCATTAGTCCCGTGCAGCCTGAAGACGACGCGGATTACTACTGCTCTGTTGGC
TACGGCTTTAGTCCC**TAG**GGGTGGGGTGTGAGATGGGTGCCTCCCCTCTGCCTCCCATTCT
GCCCCTGACCTTGGGTCCCTTTTAACTTTCTCTGAGCCTTGCTTCCCCTCTGTAAAATGGG
TTAATAATATTCAACATGTCAACAAC

FIGURE 72

MVGTKAWVFSFLVLEVTSVLGRQTMLTQSVRRVQPGKKNPSIFAKPADTLESPGEWTTWFNI
DYPGGKG DYERLDAIRFYYGDRVCARPLRLEARTTDWTPAGSTGQVVHGSPREGFWCLNREQ
RPGQNC SNYTVRFLCPPGSLRRDTERIWSPWSPWSKCSAACGQTGVQTRTRICLAEMVSLCS
EASEEGQHCMGQDCTACDLTCPMGQVNADCDACMCQDFMLHGAVSLPGGAPASGAAYLLTK
TPKLLTQTDSDGRFRIPGLCPDGKSILKITKVKFAPIVL TMPKTS LKAATIKAEFVRAETPY
MVMNPETKARRAGQSVSLCCKATGKPRPDKYFWYHNDTLLDPSLYKHESKLVLRLKLQQHQAG
EYFCKAQSDAGAVKSKVAQLIVTASDETPCNPVPESYLIRLPHDCFQ NATNSFYDVGRCPV
KTCAGQQDNGIRCRDAVQNC CGISKTEEREIQCSGYTLPTKVAKESCQRCTETRSIVRGRV
SAADNGEPMRFGHVYMGNSRVSM TGYKGTFTLHVPQDTERLVLT FVDRLQKFVN TTKVLPFN
KKGSAVFHEIKMLRRKEPITLEAMETNIIPLGEVVGEDPMAELEIPSR SFYRQNGEPYIGKV
KASVTFLDPRNISTATAAQTDLNF INDEGDTFPLRTYGMFSVDFRDEVTSEPLNAGKV KVHL
DSTQVKMPEHISTVKLWSLNPDTGLWEEEGDFKFENQRRNKREDRTFLVGNLEIRERRLFNL
DVPESRRCFVKVRAYR SERFLPSEQIQGVVISVINLEPRTGFLSNPRAWGRFDSVITGPNGA
CVPAFCDDQSPDAYSAYVLASLAGEELQAVESSPKFNPNAIGVPQPYLNKLN YRRTD HEDPR
VKKTAFQISMAKPRPNSAEESNGPIYAFENLRACEEAPPSAAHFRFYQIEGDRYDYN TVPFN
EDDPMSWTE DYLAWWPKPMEFRACYIKVKIVGPLEVNVR SRNMGGTHRRTVGKLYGIRDVRS
TRDRDQPNVSAACLEFKCSGMLYDQDRVDRTL VKVIPQGSRRASVN PMLHEYLVNHLPLAV
NNDTSEYTMLAPLDPLGHNYGIYTVTDQDPRTAKEIALGRCFDGTSDGSSRIMKSNVGVALT
FNCVERQVGRQSAFQYLQSTPAQSPAAGTVQGRVPSRRQQRASRGGQRQGGVVASLRFPRVA
QQPLIN

FIGURE 75

CCCACGCGTCCGGGTGACCTGGGCCGAGCCCTCCCGGTCGGCTAAGATTGCTGAGGAGGCGG
CGGGTAGCTGGCAGGCGCCGACTTCCGAAGGCCGCGTCCGGGCGAGGTGTCTCATGACTT
CTCTTGTGGACC**ATG**TCCGTGATCTTTTTTGCCTGCGTGGTACGGGTAAGGGATGGACTGCC
CCTCTCAGCCTCTACTGATTTTTACCACACCCAAGATTTTTTGAATGGAGGAGACGGCTCA
AGAGTTTAGCCTTGCAGCTGGCCCAGTATCCAGGTCGAGGTTCTGCAGAAGGTTGTGACTTT
AGTATACATTTTTCTTCTTTCGGGGACGTGGCCTGCATGGCTATCTGCTCCTGCCAGTGTCC
AGCAGCCATGGCCTTCTGCTTCCTGGAGACCCTGTGGTGGGAATTCACAGCTTCCTATGACA
CTACCTGCATTGGCCTAGCCTCCAGGCCATACGCTTTTCTTGAGTTTGACAGCATCATTAG
AAAGTGAAGTGGCATTTTAACTATGTAAGTTCCTCTCAGATGGAGTGCAGCTTGAAAAAAT
TCAGGAGGAGCTCAAGTTGCAGCCTCCAGCGGTTCTCACTCTGGAGGACACAGATGTGGCAA
ATGGGGTGATGAATGGTCACACACCGATGCACTTGGAGCCTGCTCCTAATTTCCGAATGGAA
CCAGTGACAGCCCTGGGTATCCTCTCCCTCATTCTCAACATCATGTGTGCTGCCCTGAATCT
CATTCGAGGAGTTCACCTTGCAGAACATTCTTTACAGGATCCAAGGAGCTGGTTCTGCTGGT
TGGACCAAACCTCG**TGA**GCCAGCCACCCCTGACCCAAATGAGGAGAGCTCTGATTCTCCCAT
CCGGGAGCAGTGATGTCAAACCTTCTGCTGCTGGGGAAATCTCATCAGCAGGGAGCCTGTGGA
AAAGGGCATGTCAGTGAAATCTGGGAATGGCTGGATTTCGGAAACATCTGCCCATGTGTATTG
ATGGCAGAGCTGTTGCCACAAGCGCCTTTTATTTAGGGTAAAATTAACAAATCCATTCTAT
TCCTCTGACCCATGCTTAGTACATATGACCTTTAACCCTTACATTTATATGATTCTGGGGTT
GCTTCAGAAGTGTTATTTTCATGAATCATTCATATGATTTGATCCCCAGGATTCTATTTTGT
TTAATGGGCTTTTCTACTAAAAGCATAAAATACTGAGGCTGATTTAGTCAGGGCAAACCAT
TTACTTTACATATTCGTTTTCAATACTTGCTGTTTCATGTTACACAAGCTTCTTACGGTTTTTC
TTGTAACAATAAATATTTTGAGTAAATAATGGGTACATTTTAACAAACTCAGTAGTACAACC
TAAACTTGATAAAAAGTGTGTAAAAATGTATAGCCATTTATATCCTATGTATAAATTAATG
AGGTGGCTTCAGAAATGGCAGAATAAATCTAAAGTGTATTATAAAAAAAAAAAAAAAAAAAAA
AAAAG

FIGURE 78

CTCAGCGGCGCTTCCTCGTAGCGAGCCTAGTGGCGGGTGTTTGCATTGAAACGTGAGCGCGA
 CCGGACCTTAAAGAGTGGGGAGCAAAGGGAGGACAGAGCCCTTTAAAACGAGGCGGGTGGTG
 CCTGCCCCCTTTAAGGGCGGGGCGTCCGGACGACTGTATCTGAGCCCCAGACTGCCCCGAGTT
 TCTGTGCGCAGGCTGCGAGGAAAGGCCCTAGGCTGGGTCTGGGTGCTTGGCGGCGGCGGCTT
 CCTCCCCGCTCGTCCTCCCCGGGCCAGAGGCACCTCGGCTTCAGTCATGCTGAGCAGAGTA
TGGGAAGCACCTGACTACGAAGTGCTATCCGTGCGAGAACAGCTATTCCACGAGAGGATCCGC
 GAGTGTATTATATCAACACTTCTGTTTGAACACTGTACATCCTCTGCCACATCTTCCTGAC
 CCGCTTCAAGAAGCCTGCTGAGTTCACCACAGTGGATGATGAAGATGCCACCGTCAACAAGA
 TTGCGCTCGAGCTGTGCACCTTTACCCTGGCAATTGCCCTGGGTGCTGTCTGCTCCTGCCC
 TTCTCCATCATCAGCAATGAGGTGCTGCTCTCCCTGCCTCGGAACACTACTACATCCAGTGGCT
 CAACGGCTCCCTCATCCATGGCCTCTGGAACCTTGTTTTCTCTTCCCCAACCTGTCCCTCA
 TCTTCCTCATGCCCTTTGCATATTTCTTCACTGAGTCTGAGGGCTTTGCTGGCTCCAGAAAG
 GGTGTCTCTGGGCCGGGTCTATGAGACAGTGGTGATGTTGATGCTCCTCACTCTGCTGGTGCT
 AGGTATGGTGTGGGTGGCATCAGCCATTGTGGACAAGAACAAGGCCAACAGAGAGTCACTCT
 ATGACTTTTGGGAGTACTATCTCCCCTACCTCTACTCATGCATCTCCTTCCTTGGGGTTCTG
 CTGCTCCTGGTGTGTACTCCACTGGGTCTCGCCCGCATGTTCTCCGTCACTGGGAAGCTGCT
 AGTCAAGCCCCGGCTGCTGGAAGACCTGGAGGAGCAGCTGTACTGCTCAGCCTTTGAGGAGG
 CAGCCCTGACCCGAGGATCTGTAATCCTACTTCCTGCTGGCTGCCTTTAGACATGGAGCTG
 CTACACAGACAGGTCCTGGCTCTGCAGACACAGAGGGTCCTGCTGGAGAAGAGGCGGAAGGC
 TTCAGCCTGGCAACGGAACCTGGGCTACCCCTGGCTATGCTGTGCTTGCTGGTGCTGACGG
 GCCTGTCTGTGCTCATTGTGGCCATCCACATCCTGGAGCTGCTCATCGATGAGGCTGCCATG
 CCGGAGGCATGCAGGGTACCTCCTTAGGCCAGGTCTCCTTCTCCAAGCTGGGCTCCTTTGG
 TGCCGTCATTACAGTTGTACTCATCTTTTACCTAATGGTGTCTCAGTTGTGGGCTTCTATA
 GCTCTCCACTCTTCCGGAGCCTGCGGCCAGATGGCACGACACTGCCATGACGCAGATAATT
 GGGAACCTGTGTCTGTCTCCTGGTCCTAAGCTCAGCACTTCCTGTCTTCTCTCGAACCCTGGG
 GCTCACTCGCTTTGACCTGCTGGGTGACTTTGGACGCTTCAACTGGCTGGGCAATTTCTACA
 TTGTGTTCTCTACAACGCAGCCTTTGCAGGCCTCACCACACTCTGTCTGGTGAAGACCTTC
 ACTGCAGCTGTGCGGGCAGAGCTGATCCGGGCCTTTGGGCTGGACAGACTGCCGCTGCCCGT
 CTCCGGTTTCCCCCAGGCATCTAGGAAGACCCAGCACCAGTGAACCTCCAGCTGGGGGTGGGA
 AGGAAAAAACTGGACACTGCCATCTGCTGCCTAGGCCTGGAGGGAAGCCCAAGGCTACTTGG
 ACCTCAGGACCTGGAATCTGAGAGGGTGGGTGGCAGAGGGGAGCAGAGCCATCTGCACTATT
 GCATAATCTGAGCCAGAGTTTGGGACCAGGACCTCCTGCTTTTCCATACTTAACTGTGGCCT
 CAGCATGGGGTAGGGCTGGGTGACTGGGTCTAGCCCCTGATCCCAAATCTGTTTACACATCA
 ATCTGCCTCACTGCTGTTCTGGGCCATCCCCATAGCCATGTTTACATGATTGATGTGCAAT
 AGGGTGGGGTAGGGGCAGGGAAGGACTGGGCCAGGGCAGGCTCGGGAGATAGATTGTCTCC
 CTGCTCTGCCCCAGCAGAGCCTAAGCACTGTGCTATCCTGGAGGGGCTTTGGACCACCTG
 AAAGACCAAGGGGATAGGGAGGAGGAGGCTTCAGCCATCAGCAATAAAGTTGATCCCAGGGA
 AAAAAA

FIGURE 82

GATGTGCTCCTTGGAGCTGGTGTGCAGTGTCTGACTGTAAGATCAAGTCCAAACCTGTTTT
GGAATTGAGGAACTTCTCTTTTGATCTCAGCCCTTGGTGGTCCAGGTCTTC**CATG**CTGCTGT
GGGTGATATTACTGGTCCTGGCTCCTGTCAGTGGACAGTTTGCAAGGACACCCAGGCCCAT
ATTTTCCTCCAGCCTCCATGGACCACAGTCTTCCAAGGAGAGAGAGTGACCCTCACTTGCAA
GGGATTTTCGCTTCTACTCACCACAGAAAACAAAATGGTACCATCGGTACCTTGGGAAAGAAA
TACTAAGAGAAACCCAGACAATATCCTTGAGGTTTCAGGAATCTGGAGAGTACAGATGCCAG
GCCCAGGGCTCCCCCTCTCAGTAGCCCTGTGCACTTGGATTTTTCTTCAGAGATGGGATTTCC
TCATGCTGCCCAGGCTAATGTTGAACTCCTGGGCTCAAGTGATCTGCTCACCT**TAG**GCCTCTC
AAAGCGCTGGGATTACAGCTTCGCTGATCCTGCAAGCTCCACTTTCTGTGTTTGAAGGAGAC
TCTGTGGTTCTGAGGTGCCGGGCAAAGGCGGAAGTAACACTGAATAATACTATTTACAAGAA
TGATAATGTCCTGGCATTCTTAATAAAAGAACTGACTTCCAAAAAAAAAAAAAAAAAAAAA
AAA

83/330

FIGURE 83

MLLWVILLVLAPVSGQFARTPRPIIFLQPPWTTVFQGERVTLTCKGFRFYSPQGTKWYHRYL
GKEILRETPDNILEVQESGEYRCQAQGSPLSSPVHLDFSSEMGFPHAAQANVELLGSSDLLT

0999727 14994
105111 2229990

FIGURE 85

MAPQNLSTFCLLLLYLIGAVIAGRDFYKILGVPRSASIKDIKKAYRKLALQLHPDRNPDDPQ
AQEKFQDLGAAYEVLSDSEKQYDQTYGEEGLKDGHQSSHGDI FSHFFGDFGFMFGGTPRQQ
DRNIPRGSDIIVDLEVTLEEYAGNFVEVVRNKPVARQAPGKRKCNCRQEMRTTQLGPGRFQ
MTQEVVCDECPNVKLVNEERTLEVEIEPGVRDGMETPFGEGEAPHVDGEPGDLRFRIKVVKH
PIFERRGDDLYTNVTISLVESLVGFEMDITHLDGHKVHISRDKITRPGAKLWKKGEGLPNFD
NNNIKGSIIITFDVDFPKEQLTEEAREGIKQLLKQGSVQKVYNGLQGY

Important features:**Signal peptide:**

amino acids 1-22

Cell attachment sequence.

amino acids 254-257

Nt-dnaJ domain signature.

amino acids 67-87

Homologous region to Nt-dnaJ domain proteins.

amino acids 26-58

N-glycosylation site.

amino acids 5-9, 261-265

Tyrosine kinase phosphorylation site.

amino acids 253-260

N-myristoylation site.

amino acids 18-24, 31-37, 93-99, 215-221

Amidation site.

amino acids 164-168

FIGURE 86

TGGGACCAGGGAACCCCGGGCCCCCGGTGGAGNGCCTAACAGGCCGGTGGNTGCGACCGAA
GCGGCGGGCGGAGGAGGTTTTGAGGATTTTGGAAACAGGACCCGGACAGAGGAACCATGGTT
CCGCAGAACNTGAGCACNTTTTGCCTGTTGNTGNTATACTTCATCGGGGCGGTGATTGCCGG
ACGAGATTTNTATAAGATTTTGGGGTGCCCTNGAAGTGCCTTNTATAAAGGATATTAAAAAGG
CCTATAGGAAACTAGCCCTGCAGNTTATCCCGACCGGAACCCTGATGATCCACAAGCCCAG
GAGAAATTCAGGATTTGGGTGCTGCTTATGAGGTTNTGTCAGATAGTGAGAAACGGAAACA
GTACGATAATTATGGTGAAGAAGGATTAAGATGGTNATCAGAGCTCCCATGGAGACATTT
TTTCACACTTNTTTGGGGATTTTGGTTTCATGTTTGGAGGAACCCCTNGTCAGCAAGACAGA
AATATTCCAAGAG

FIGURE 87

GGCACGAGGCGGCGGGGCAGTCGCGGGATGCGCCGGGAGCCACAGCCTGAGGCCCTCAGGT
CTCTGCAGGTGTCGTGGAGGAACCTAGCACCTGCCATCCTCTTCCCCAATTTGCCACTTCCA
GCAGCTTTAGCCCATGAGGAGGATGTGACCGGGACTGAGTCAGGAGCCCTCTGGAAGC**ATGG**
AGACTGTGGTGATTGTTGCCATAGGTGTGCTGGCCACCATCTTTCTGGCTTCGTTTGCAGCC
TTGGTGCTGGTTTGCAGGCAGCGCTACTGCCGGCCGCGAGACCTGCTGCAGCGCTATGATTC
TAAGCCCATTGTGGACCTCATTGGTGCCATGGAGACCCAGTCTGAGCCCTCTGAGTTAGAAC
TGGACGATGTCGTTATCACCAACCCCCACATTGAGGCCATTCTGGAGAATGAAGACTGGATC
GAAGATGCCTCGGGTCTCATGTCCCACTGCATTGCCATCTTGAAGATTTGTCACACTCTGAC
AGAGAAGCTTGTTGCCATGACAATGGGCTCTGGGGCCAAGATGAAGACTTCAGCCAGTGTCA
GCGACATCATTGTGGTGGCCAAGCGGATCAGCCCCAGGGTGGATGATGTTGTGAAGTCGATG
TACCCTCCGTTGGACCCCAAACCTCCTGGACGCACGGACGACTGCCCTGCTCCTGTCTGTGAC
TCACCTGGTGCTGGTGACAAGGAATGCCTGCCATCTGACGGGAGGCCTGGACTGGATTGACC
AGTCTCTGTGCGGCTGCTGAGGAGCATTTGGAAGTCCTTCGAGAAGCAGCCCTAGCTTCTGAG
CCAGATAAAGGCCTCCCAGGCCCTGAAGGCTTCCTGCAGGAGCAGTCTGCAATTT**TAGT**GCCT
ACAGGCCAGCAGCTAGCCATGAAGGCCCTGCCGCCATCCCTGGATGGCTCAGCTTAGCCTT
CTACTTTTTCTATAGAGTTAGTTGTTCTCCACGGCTGGAGAGTTCAGCTGTGTGTGCATAG
TAAAGCAGGAGATCCCCGTCAGTTTATGCCTCTTTTGCAGTTGCAAACGTGGCTGGTGAGT
GGCAGTCTAATACTACAGTTAGGGGAGATGCCATTCACTCTCTGCAAGAGGAGTATTGAAAA
CTGGTGGACTGTCAGCTTTATTTAGCTCACCTAGTGTTTTCAAGAAAATTGAGCCACCGTCT
AAGAAATCAAGAGGTTTCACATTAAAATTAGAATTTCTGGCCTCTCTCGATCGGTCAGAATG
TGTGGCAATTCTGATCTGCATTTTTCAGAAGAGGACAATCAATTGAAACTAAGTAGGGGTTTC
TTCTTTTGGCAAGACTTGTACTCTCTCACCTGGCCTGTTTCATTTATTTGTATTATCTGCCT
GGTCCCTGAGGCGTCTGGGTCTCTCCTCTCCCTTGCCAGGTTTGGGTTTGAAGCTGAGGAACT
ACAAAGTTGATGATTTCTTTTTTATCTTTATGCCTGCAATTTTACCTAGCTACCACTAGGTG
GATAGTAAATTTATACTTATGTTTTCCCTCAAAAAAAAAAAAAA

FIGURE 89

GCTTCATTTCTCCCGACTCAGCTTCCCACCCTGGGCTTTCCGAGGTGCTTTCGCCGCTGTCC
CCACCACTGCAGCC**ATG**ATCTCCTTAACGGACACGCAGAAAATTGGAATGGGATTAACAGGA
TTTGGAGTGTTTTTCCTGTTCTTTGGAATGATTCTCTTTTTTGACAAAGCACTACTGGCTAT
TGAAATGTTTTATTTGTAGCCGGCTTGGCTTTTGTAATTGGTTTAGAAAGAACATTTCAGAT
TCTTCTTCCAAAACATAAAATGAAAGCTACAGGTTTTTTTTCTGGGTGGTGTATTTGTAGTC
CTTATTGGTTGGCCTTTGATAGGCATGATCTTCGAAATTTATGGATTTTTTCTCTTGTTTCAG
GGGCTTCTTTCCTGTCGTTGTTGGCTTTATTAGAAGAGTGCCAGTCCTTGGATCCCTCCTAAAT
TTACCTGGAATTAGATCATTGTAGATAAAGTTGGAGAAAGCAACAATATGGTAT**TAA**CAACA
AGTGAATTTGAAGACTCATTTAAAATATTGTGTTATTTATAAAGTCATTTGAAGAATATTCA
GCACAAAATTAAATTACATGAAATAGCTTGTAATGTTCTTTACAGGAGTTTAAAACGTATAG
CCTACAAAGTACCAGCAGCAAATTAGCAAAGAAGCAGTGAAAACAGGCTTCTACTCAAGTGA
ACTAAGAAGAAGTCAGCAAGCAAAGTGAAGAGAGGTGAAATCCATGTTAATGATGCTTAAGAA
ACTCTTGAAGGCTATTTGTGTTGTTTTTCCACAATGTGCGAAACTCAGCCATCCTTAGAGAA
CTGTGGTGCCTGTTTCTTTTCTTTTTATTTTGAAGGCTCAGGAGCATCCATAGGCATTTGCT
TTTTAGAAGTGTCCTACTGCAATGGCAAAAATATTTCCAGTTGCACTGTATCTCTGGAAGTGA
TGCATGAATTCGATTGGATTGTGTCATTTTAAAGTATTAAAACCAAGGAAACCCCAATTTTG
ATGTATGGATTACTTTTTTTTTGNGCNCAGGGCC

FIGURE 91

GAAGACGTGGCGGCTCTCGCCTGGGCTGTTTCCCGGCTTCATTTCTCCCGACTCAGCTTCCC
ACCNTGGGCTTTCCGAGGTGCTTTCGCCGCTGTCCCCACCACTGCAGCCATGATCTCCTTAA
CGGACACGCAGAAAATTGGAATGGGATTAACCGGATTTGGAGTGTTTTCTGTCTTTGGA
ATGATTCTCTTTTTTGACAAAGCACTACTGGCTATTGGAAATGTTTTATTTGTAGCCGGCTT
GGCTTTTGTAATTGGTTTAGAAAGAACATTCAGATTCTTCTTCCAAAAACATAAAATGAAAG
CTACAGGTTTTTTTCTGGGTGGTGTATTTGTAGTCCTTATTGGTTGGCCTTTGATAGGCATG
ATCTTCGAAATTTATGGATTTTTTCTCTTGTTT

93/330

FIGURE 93

MDLAGLLKSQFLCHLVFCYVFIA SGLIINTIQLF TLLLWPINKQLFRKINCRLSYCISSQLV
MLLEWWSGTECTIFTDPRAYLKYGKENAIVVLN HKFEIDFLCGWSLSE RFGLLGGSKVLAKK
ELAYVPIIGWMWYFTEMVFCSRKWEQDRKTVATSLQH LRDYPEKYFFLIHCEGTRFTEKKHE
ISMQVARAKGLPRLKHHLLPRTKGFAITVRSLRNVVSAVYDCTLNFRNNENPTLLGV LNKKK
YHADLYVRRIPLEDIPEDDDECSAWLHKLYQEKDAFQEEYYRTGTFPETPMVPPRRPWTLVN
WLFWASLVLYPFFQFLVSMIRSGSSLTLASFILVFFVASVGVRWMIGVTEIDKGSAYGNSDS
KQKLND

FOOT 226350

FIGURE 94

CTGAGGCGGCGGTAGC**ATG**GAGGGGGAGAGTACGTCGGCGGTGCTCTCGGGCTTTGTGCTCG
 GCGCACTCGCTTTCCAGCACCTCAACACGGACTCGGACACGGAAGGTTTTCTTCTTGGGGAA
 GTAAAAGGTGAAGCCAAGAACAGCATTACTGATTCCCAAATGGATGATGTTGAAGTTGTTTA
 TACAATTGACATTCAGAAATATATTCCATGCTATCAGCTTTTTAGCTTTTATAATTCTTCAG
 GCGAAGTAAATGAGCAAGCACTGAAGAAAATATTATCAAATGTCAAAAAGAATGTGGTAGGT
 TGGTACAAATTCCGTCGTCATTCAGATCAGATCATGACGTTTAGAGAGAGGCTGCTTCACAA
 AACTTGCAGGAGCATTTTTCAAACCAAGACCTTGTTTTTCTGCTATTAACACCAAGTATAA
 TAACAGAAAGCTGCTCTACTCATCGACTGGAACATTCCTTATATAAACCTCAAAAAGGACTT
 TTTCACAGGGTACCTTTAGTGGTTGCCAATCTGGGCATGTCTGAACAACCTGGGTATATAAAC
 TGTATCAGGTTCTGTATGTCCACTGGTTTTAGCCGAGCAGTACAAACACACAGCTCTAAAT
 TTTTTGAAGAAGATGGATCCTTAAAGGAGGTACATAAGATAAATGAAATGTATGCTTCATTA
 CAAGAGGAATTAAAGAGTATATGCAAAAAAGTGGAAGACAGTGAACAAGCAGTAGATAAACT
 AGTAAAGGATGTAAACAGATTAAAACGAGAAATTGAGAAAAGGAGAGGAGCACAGATTCAGG
 CAGCAAGAGAGAAGAACATCCAAAAAGACCCTCAGGAGAACATTTTTCTTTGTCAGGCATTA
 CGGACCTTTTTTCCAAATTCTGAATTTCTTCATTCATGTGTTATGTCTTTAAAAAATAGACA
 TGTTTTCTAAAAGTAGCTGTAACACACCACCATCTCGATGTAGTAGACAATCTGACCTTAA
 TGGTAGAACACACTGACATTCCTGAAGCTAGTCCAGCTAGTACACCACAAATCATTAAGCAT
 AAAGCCTTAGACTTAGATGACAGATGGCAATTCAAGAGATCTCGGTTGTTAGATACACAAGA
 CAAACGATCTAAAGCAAATACTGGTAGTAGTAACCAAGATAAAGCATCCAAAATGAGCAGCC
 CAGAAACAGATGAAGAAATTGAAAAGATGAAGGGTTTTGGTGAATATTCACGGTCTCCTACA
 TTT**TGA**TCCTTTTAACCTTACAAGGAGATTTTTTTATTTGGCTGATGGGTAAAGCCAAACAT
 TTCTATTGTTTTTACTATGTTGAGCTACTTGCAGTAAGTTCATTTGTTTTTACTATGTTTAC
 CTGTTTGCAGTAATACACAGATAACTCTTAGTGCATTTACTTCACAAAGTACTTTTTCAAAC
 ATCAGATGCTTTTATTTCCAAACCTTTTTTTTACCTTTCACTAAGTTGTTGAGGGGAAGGCT
 TACACAGACACATTCCTTTAGAATTGGAAAAGTGAGACCAGGCACAGTGGCTCACACCTGTAA
 TCCCAGCACTTAGGGAAGACAAGTCAGGAGGATTGATTGAAGCTAGGAGTTAGAGACCAGCC
 TGGGCAACGTATTGAGACCATGTCTATTAAAAAATAAAATGGAAAAGCAAGAATAGCCTTAT
 TTCAAATATGGAAGAAATTTATATGAAAATTTATCTGAGTCATTAAAATTCTCCTTAAG
 TGATACTTTTTTAGAAGTACATTATGGCTAGAGTTGCCAGATAAAATGCTGGATATCATGCA
 ATAAATTTGCAAAACATCATCTAAAATTTAAAAAATAAAAAAAAAAAAAAAAAAAAAA

FIGURE 95

MEGESTSAVLSGFVLGALAFQHLNTDSDTEGFLLGEVKGEAKNSITDSQMDDVEVVYTIDIQ
KYIPCYQLFSFYNSSGEVNEQALKKILSNVKKNVVGWYKFRRHSDQIMTFRERLLHKNLQEH
FSNQDLVFLLLTPSIITESCSTHRLEHSLYKPQKGLFHRVPLVVANLGMSEQLGYKTVSGSC
MSTGFSRAVQTHSSKFFEEEDGSLKEVHKINEMYASLQEELKSICKKVEDSEQAVDKLVKDVN
RLKREIEKRRGAQIQAAAREKNIQKDPQENIFLCQALRTFFPNSEFLHSCVMSLKNRHVSKSS
CNYNHHLDVVDNLTLMVEHTDIPEASPASTPQIIKHKALDLDDRWFKRSRLLDTDQDKRSKA
NTGSSNQDKASKMSSPETDEEIEKMKGFGEYSRSPTF

0999727 44904
10611 2228850

FIGURE 96

GGCACAGCCGCGCGGCGGAGGGCAGAGTCAGCCGAGCCGAGTCCAGCCGGACGAGCGGACCAGCGCAGGGCAGC
 CCAAGCAGCGCGCAGCGAACGCCCGCCGCCACACCCTCTGCGGTCCCCGCGGCGCCTGCCACCCTTCCCT
 CCTTCCCCGCGTCCCCGCCCTCGCCGGCCAGTCAGCTTGCCGGGTTGCTGCCCCGCGAAACCCCGAGGTACCA
 GCGCGCGCCTCTGCTTCCCTGGGCGCGCGCGCCCTCCACGCCCTCCTTCTCCCTGGCCCGCGCCTGGCACC
 GGGGACCGTTGCCTGACGCGAGGCCAGCTCTACTTTTGGCCCCGCTCTCCTCCGCTGCTCGCCTCTTCCAC
 CAACTCCAACTCCTTCTCCCTCCAGCTCCACTCGCTAGTCCCCGACTCCGCCAGCCCTCGGCCCGCTGCCGTAG
 CGCCGCTTCCCGTCCGGTCCCAAAGGTGGGAACGCGTCCGCCCGGCCCGCACC**ATG**GCACGGTTCCGGCTTGCC
 CGCGCTTCTCTGCACCTGGCAGTGCTCAGCGCCGCGCTGCTGGCTGCCGAGCTCAAGTCGAAAAGTTGCTCGG
 AAGTGCAGAGCTCTTTACGTGTCCAAAGGCTTCAACAAGAACGATGCCCCCTCCACGAGATCAACGGTGATCAT
 TTGAAGATCTGTCCCCAGGGTTCTACCTGCTGCTCTCAAGAGATGGAGGAGAAGTACAGCCTGCAAAGTAAAGA
 TGATTTCAAAGTGTGGTCAGCGAACAGTGCAATCATTTGCAAGCTGTCTTTGCTTACGTTACAAGAAGTTTG
 ATGAATTTCTCAAAGAAGTACTTGAAAATGCAGAGAAATCCCTGAATGATATGTTTGTGAAGACATATGGCCAT
 TTATACATGCAAAATTTCTGAGCTATTTAAAGATCTCTTCGTAGAGTTGAAACGTTACTACGTGGTGGGAAATGT
 GAACCTGGAAGAAATGCTAAATGACTTCTGGGCTCGCCTCCTGGAGCGGATGTTCCGCTGTTGAAGTCCCAGT
 ACCACTTTACAGATGAGTATCTGGAATGTGTGAGCAAGTATACGGAGCAGCTGAAGCCCTTCGAGATGTCCCT
 CGCAAATTTGAAGCTCCAGGTACTCGTGCTTTTGTAGCAGCCCGTACTTTTCGCTCAAGGCTTAGCGGTTGCGGG
 AGATGTCGTGAGCAAGGTCTCCGTGGTAAACCCACAGCCAGTGTACCCATGCCCTGTTGAAGATGATCTACT
 GCTCCCACTGCCGGGTCTCGTGACTGTGAAGCCATGTTACAACACTGCTCAAACATCATGAGAGGCTGTTTG
 GCCAACCAAGGGGATCTCGATTTTGAATGGAACAATTTATAGATGCTATGCTGATGGTGGCAGAGAGGCTAGA
 GGGTCCTTTCAACATTGAATCGGTATGGATCCCATCGATGTGAAGATTTCTGATGCTATTATGAACATGCAGG
 ATAATAGTGTCAAGTGTCTCAGAAGGTTTCCAGGGATGTGGACCCCCCAAGCCCTCCAGCTGGACGAATT
 TCTCGTTCCATCTCTGAAAGTGCCTTCAGTGCTCGCTTCAGACCACATCACCCCGAGGAACGCCCAACCACAGC
 AGCTGGCACTAGTTTGGACCGACTGGTTACTGATGTCAAGGAGAACTGAAACAGGCCAAGAAATTTCTGGTCCCT
 CCCTTCCGAGCAACGTTTGAACGATGAGAGGATGGCTGCAGGAAACGGCAATGAGGATGACTGTTGGAATGGG
 AAAGGCAAAAGCAGGTACCTGTTTGCAGTGACAGGAAATGGATTAGCCAACCAGGGCAACAACCCAGAGGTCCA
 GGTGACACCAGCAAACCAGACATACTGATCCTTCGTCAAATCATGGCTCTTCGAGTGATGACCAGCAAGATGA
 AGAATGCATACAATGGGAACACGTGGACTTCTTTGATATCAGTGATGAAAGTAGTGGAGAAGGAAGTGGAAGT
 GGCTGTGAGTATCAGCAGTGCCCTTCAGAGTTTACTACAATGCCACTGACCATGCTGGGAAGAGTGCCAAATGA
 GAAAGCCGACAGTGCTGGTGTCCGTCTGGGGCACAGGCCTACCTCCTCACTGTCTTCTGCATCTTGTTCCTGG
 TTATGCAGAGAGAGTGGAGAT**TAA**TTCTCAAACCTCTGAGAAAAAGTGTTTCATCAAAAAGTTAAAAGGCACCAAGTT
 ATCACTTTTCTACCATCCTAGTGACTTTGCTTTTTTAAATGAATGGACAACAATGTACAGTTTTTACTATGTGGC
 CACTGGTTTTAAGAAGTGCTGACTTTGTTTTCTCATTAGTTTTGGGAGGAAAAGGGACTGTGCATTGAGTTGGT
 TCCTGCTCCCCCAAACCATGTTAAACGTGGCTAACAGTGTAGGTACAGAACTATAGTTAGTTGTGCATTTGTGA
 TTTTATCACTCTATTATTTGTTTGTATGTTTTTTCTCATTTCGTTTGTGGGTTTTTTTTTCCAACGTGATCT
 CGCCTTGTCTTACAAGCAAACAGGGTCCCTTCTTGGCACGTAACATGTACGTATTTCTGAAATATTAAATA
 GCTGTACAGAAGCAGGTTTTATTTATCATGTTATCTTATTTAAAGAAAAAGCCCAAAAAGC

FIGURE 98

CTCGCCCTCAAATGGGAACGCTGGCCTGGGACTAAAGCATAGACCACCAGGCTGAGTATCCT
GACCTGAGTCATCCCCAGGGATCAGGAGCCTCCAGCAGGGAACCTTCCATTATATTCTTCAA
GCAACTTACAGCTGCACCGACAGTTGCG**ATG**AAAGTTCTAATCTCTTCCCTCCTCCTGTTGC
TGCCACTAATGCTGATGTCCATGGTCTCTAGCAGCCTGAATCCAGGGGTCGCCAGAGGCCAC
AGGGACCGAGGCCAGGCTTCTAGGAGATGGCTCCAGGAAGGCGGCCAAGAATGTGAGTGCAA
AGATTGGTTCCTGAGAGCCCCGAGAAGAAAATTTCATGACAGTGTCTGGGCTGCCAAAGAAGC
AGTGCCCCTGTGATCATTTCAAGGGCAATGTGAAGAAAACAAGACACCAAAGGCACCACAGA
AAGCCAAACAAGCATTCCAGAGCCTGCCAGCAATTTCTCAAACAATGTCAGCTAAGAAGCTT
TGCTCTGCCTTTG**TAG**GAGCTCTGAGCGCCCACTCTTCCAATTAAACATTCTCAGCCAAGAA
GACAGTGAGCACACCTACCAGACACTCTTCTTCTCCACCTCACTCTCCCACTGTACCCACC
CCTAAATCATTCCAGTGCTCTCAAAAAGCATGTTTTTCAAGATCATTTTGTTTGTTGCTCTC
TCTAGTGTCTTCTTCTCTCGTCAGTCTTAGCCTGTGCCCTCCCCTTACCCAGGCTTAGGCTT
AATTACCTGAAAGATTCCAGGAACTGTAGCTTCCTAGCTAGTGTCAATTTAACCTTAAATGC
AATCAGGAAAGTAGCAAACAGAAGTCAATAAATATTTTTAAATGTCAAAAAAAAAAAAAAAAAA

99/330

FIGURE 99

MKVLISLLLLLPLMLMSMVSSSLNPGVARGHRDRGQASRRWLQEGGQECECKDWFLRAPRR
KFMTVSGLPKKQCPDHFKGNVKKTRHQRRHRKPNKHSRACQQLKQCQLRSFALPL

0908727 1404
TOT 26360

100/330

FIGURE 100

AATGGCTGTCTTAGTACTTCGCCTGACAGTTGTCCTGGGACTGCTTGTCTTATTCCTGACCT
GCTATGCAGACGACAAACCAGACAAGCCAGACGACAAGCCAGACGACTCGGGCAAAGACCCA
AAGCCAGACTTCCCCAAATTCCTAAGCCTCCTGGGCACAGAGATCATTGAGAATGCAGTCGA
GTTTCATCCTCCGCTCCATGTCCAGGAGCACAGGATTTATGGAATTTGATGATAATGAAGGAA
AACATTCATCAAAG**TGA**CATCCTCAGGACACACCCATGTGGCTCCTGGACAATCCAAGAGCA
GCCAAATCCTGCTTTTCCAGTTTGGCTCCACAAGTCCTCCAGGACAGAGCCCTCAAAGCAAC
TCCCAACGAGTTCTCAGGATTCAGGCTCTGGCTTCAACCAAACAGAACTCATTTTGAACACC
CTGACTGCATTTTTGCTTTTAGAAAGTTAGAATAAATATGGCGCTTTGGGATCACATAGTTG
ATGGAGAGGAAA

0999727-11001

FIGURE 102

GGACGCCAGCGCCTGCAGAGGCTGAGCAGGGAAAAAGCCAGTGCCCCAGCGGAAGCACAGCT
CAGAGCTGGTCTGCCATGGACATCCTGGTCCCACCTCCTGCAGCTGCTGGTGCTGCTTCTTAC
CCTGCCCCCTGCACCTCATGGCTCTGCTGGGCTGCTGGCAGCCCCCTGTGCAAAGCTACTTCC
CCTACCTGATGGCCGTGCTGACTCCCAAGAGCAACCGCAAGATGGAGAGCAAGAAACGGGAG
CTCTTCAGCCAGATAAAGGGGCTTACAGGAGCCTCCGGGAAAGTGGCCCTACTGGAGCTGGG
CTGCGGAACCGGAGCCAACTTTCAGTTCTACCCACCGGGCTGCAGGGTCACCTGCCTAGACC
CAAATCCCCACTTTGAGAAGTTCCTGACAAAGAGCATGGCTGAGAACAGGCACCTCCAATAT
GAGCGGTTTGTGGTGGCTCCTGGAGAGGACATGAGACAGCTGGCTGATGGCTCCATGGATGT
GGTGGTCTGCACTCTGGTGCTGTGCTCTGTGCAGAGCCCCAAGGAAGGTCTGCAGGAGGTCC
GGAGAGTACTGAGACCGGGAGGTGTGCTCTTTTTCTGGGAGCATGTGGCAGAACCATATGGA
AGCTGGGCCTTCATGTGGCAGCAAGTTTTTCGAGCCCACCTGGAAACACATTGGGGATGGCTG
CTGCCTCACCAGAGAGACCTGGAAGGATCTTGAGAACGCCCAGTTCTCCGAAATCCAAATGG
AACGACAGCCCCCTCCCTTGAAGTGGCTACCTGTTGGGCCCCACATCATGGGAAAGGCTGTC
AAACAATCTTTCCCAAGCTCCAAGGCACTCATTTGCTCCTTCCCCAGCCTCCAATTAGAACA
AGCCACCCACCAGCCTATCTATCTTCCACTGAGAGGGACCTAGCAGAATGAGAGAAGACATT
CATGTACCACCTACTAGTCCCTCTCTCCCCAACCTCTGCCAGGGCAATCTCTAACTTCAATC
CCGCCTTCGACAGTGAAAAAGCTCTACTTCTACGCTGACCCAGGGAGGAAACACTAGGACCC
TGTTGTATCCTCAACTGCAAGTTTCTGGACTAGTCTCCCAACGTTTGCCTCCCAATGTTGTC
CCTTTCCTTCGTTCCCATGGTAAAGCTCCTCTCGCTTTCCTCCTGAGGCTACACCCATGCGT
CTCTAGGAACTGGTCACAAAAGTCATGGTGCCTGCATCCCTGCCAAGCCCCCTGACCCTCT
CTCCCCACTACCACCTTCTTCTGAGCTGGGGGCACCAGGGAGAATCAGAGATGCTGGGGAT
GCCAGAGCAAGACTCAAAGAGGCAGAGGTTTTGTTCTCAAATATTTTTTAATAAATAGACGA
AACCACG

FIGURE 104

GTGGGATTTATTTGAGTGCAAGATCGTTTTCTCAGTGGTGGTGGGAAGTTGCCTCATCGCAGG
CAGATGTTGGGGCTTTGTCCGAACAGCTCCCCCTCTGCCAGCTTCTGTAGATAAGGGTTAAAA
ACTAATATTTATATGACAGAAGAAAAAGATGTCATTCCGTAAAGTAAACATCATCATCTTGG
TCCTGGCTGTTGCTCTCTTCTTACTGGTTTTGCACCATAACTTCCTCAGCTTGAGCAGTTTG
TTAAGGAATGAGGTTACAGATTCAGGAATTGTAGGGCCTCAACCTATAGACTTTGTCCCAA
TGCTCTCCGACATGCAGTAGATGGGAGACAAGAGGAGATTCTGTGGTCATCGCTGCATCTG
AAGACAGGCTTGGGGGGGCCATTGCAGCTATAAACAGCATTTCAGCACAACTCGCTCCAAT
GTGATTTTCTACATTGTTACTCTCAACAATACAGCAGACCATCTCCGGTCCTGGCTCAACAG
TGATTCCCTGAAAAGCATCAGATACAAAATTGTCAATTTTGACCCTAAACTTTTGGAAGGAA
AAGTAAAGGAGGATCCTGACCAGGGGGAATCCATGAAACCTTTAACCTTTGCAAGGTTCTAC
TTGCCAATTCTGGTTCCCAGCGCAAAGAAGGCCATATACATGGATGATGATGTAATTGTGCA
AGGTGATATTCTTGCCCTTTACAATACAGCACTGAAGCCAGGACATGCAGCTGCATTTTCAG
AAGATTGTGATTCAGCCTCTACTAAAGTTGTCATCCGTGGAGCAGGAAACCAGTACAATTAC
ATTGGCTATCTTGACTATAAAAAGGAAAGAATTCGTAAGCTTTCCATGAAAGCCAGCACTTG
CTCATTTAATCCTGGAGTTTTTGTGCAAACCTGACGGAATGGAAACGACAGAATATAACTA
ACCAACTGGAAAAATGGATGAACTCAATGTAGAAGAGGGACTGTATAGCAGAACCTGGCT
GGTAGCATCACAACACCTCCTCTGCTTATCGTATTTTATCAACAGCACTCTACCATCGATCC
TATGTGGAATGTCCGCCACCTTGGTTCCAGTGCTGGAAAACGATATTCACCTCAGTTTGTA
AGGCTGCCAAGTTACTCCATTGGAATGGACATTTGAAGCCATGGGGAAGGACTGCTTCATAT
ACTGATGTTTGGGAAAAATGGTATATTCAGACCCAACAGGCAAATTCAACCTAATCCGAAG
ATATACCGAGATCTCAAACATAAAGTGAACAGAAATTTGAACTGTAAGCAAGCATTCTCAG
GAAGTCCTGGAAGATAGCATGCATGGGAAGTAACAGTTGCTAGGCTTCAATGCCTATCGGTA
GCAAGCCATGGAAAAAGATGTGTCAGCTAGGTAAAGATGACAACTGCCCTGTCTGGCAGTC
AGCTTCCCAGACAGACTATAGACTATAAATATGTCTCCATCTGCCTTACCAAGTGTTCCTT
ACTACAATGCTGAATGACTGGAAAGAAGAACTGATATGGCTAGTTCAGCTAGCTGGTACAGA
TAATTCAAACTGCTGTTGGTTTTAATTTTGTAACCTGTGGCCTGATCTGTAAATAAACTT
ACATTTTTC

105/330

FIGURE 105

MSFRKVNIIILVLAVALFLLVLHHNFLSLSSLLRNEVTD SGIVGPQPIDFVPNALRHAVDGR
QEEIPVVIAASEDRLGGAIAAINSIQHNTSRNVIFYIVTLNNTADHLRSWLNSDSLKSIRYK
IVNFDPKLLEGKVKEDPDQGESMKPLTFARFYLPILVPSAKKAIYMDDDVIVQGDILALYNT
ALKPGHAAAFSEDCDSASTKVVIRGAGNQYNYIGYLDYKKERIRKLSMKASTCSFNPGVFVA
NLTEWKRQINITNQLEKWMKLNVEEGLYSRTLGSITTPPLLIVFYQQHSTIDPMWNVRLGS
SAGKRYSPQFVKAAKLLHWNHGLKPWGRTASYTDVWEKWYIPDPTGKFNLIRRYTEISNIK

090972741994

FIGURE 106

TGGTTTTTGCCCCATAAATTCCTCAGCTTGAGCAGTTTGTTAAGGAATGAGGTTACAGATT
CAGGAATTNTAGGNCCTCAACCTNTAGANTTTGTCCCAAATGTTCTCCGACATGCAGTAGAT
GGGAGACAAGAGGAGATTCTGTGGTCATCGCTGCATNTGAAGACAGGCTTGGGGGGGCCAT
TGCAGCTATAAACAGCATTTCAGCACAACTCGNTCCAATGTGATTTTCTACATTGTTACTC
TCAACAATACAGCAGACCATNTCCGGTCCTGGNTCAACAGTGATTCCCTGAAAAGCATCAGA
TACAAAATTGTCAATTTTGACCCTAACTTTTGGAAGGAAAAGTAAAGGAGGATCCTGACCA
GGGGGAATCCATGAAACCTTTAACCTTTGCAAGGTTCTACTTGCCAATTCTGGTTCCCAGCG
CAAAGAAGGCCATATACATGGATGATGATGTAATTGTGCAAGGTGATATTCTTGCCCTTTAC
AATACAGCACTGAAGCCAGGACATGCAGCTGCATTTTCAGAAGATTGTGATTCAGCCTCTAC
TAAAGTTGTCATCCGTGGAGCAGGAAA

106/330

FIGURE 107

CGACGCTCTAGCGGTTACCGCTGCGGGCTGGCTGGGCGTAGTGGGGCTGCGCGGCTGCCACG
GAGCTAGAGGGCAAGTGTGCTCGGCCAGCGTGCAGGGAACGCGGGCGGCCAGACAACGGGC
TGGGCTCCGGGGCCTGCGGCGCGGGCGCTGAGCTGGCAGGGCGGGTCGGGGCGCGGGCTGCA
TCCGCATCTCCTCCATCGCCTGCAGTAAGGGCGGGCGCGGCGAGCCTTTGAGGGGAACGACT
TGTCGGAGCCCTAACCAGGGGTGTCTCTGAGCCTGGTGGGATCCCCGAGCGTCACATCACT
TTCCGATCACTTCAAAGTGGTTAAAACTAATATTTATATGACAGAAGAAAAAGATGTCATT
CCGTAAAGTAAACATCATCATCTTGGTCTGGGCTGTTGCTCTCTTCTTACTGGTTTTGCAC
CATAACTTCCTCAGCTTGAGGCAGTTTGTAAAGGAATGAGGTTACAGATTCAGGAATTGTAG
GGCCTCAACCTATAGGACTTTGTCCCAAATGCTCTCCGACATGCAGTAGATGGGAGACAAGA
GGAGATTCTGTGGTCATCGCTGCATCTGAAGACAGGCTTGGGGGGGCCATTGCAGCTATAA
ACAGCATTTCAGCACAACTCGCTCCAATGTGATTTTCTACATTGTTACTCTCAACAATACA
GCAGACCATCTCCGGTCCTGGGCTCAACAGTGATTCCCTGAAAAGCATCAGATACAAAATTG
TCAATTTTGACCCTAAACTTTTGGAAGGAAAAGTAAAGGAGGATCCTGACCAGGGGGAATCC
ATGAAACCTTTAACCTTTGCAAGGTTCTACTTGCCAATTCTGGGTTCCCAGCGCAAAGAAGG
CCATATACATGGATGATGATGTAATTGTGCAAGGTGATATTCTTGCCCTTTACAATACAGCA
CTGAAGCCAGGACATGCAGCTGCATTTTTCAGAAGATTGTGATTCAGCCTCTACTAAAGTTGT
CATCCGTGGAGCAGGAAACCAGTACAATTACATTGGCTATCTTGACTATAAAAAGGAAAGAA
TTCGTAAGCTTTCCATGAAAGCCAGCACTTGCTCATTTAATCCTGGAGTTTTTGTGCAAAC
CTGACGGAATGGAAACGACAGAATATACTAACCAACTGGAAAAATGGATGAAACTCAATGT
AGAAGAGGGACTGTATAGCAGAACCCTGGCTGGTAGCATCACAAACACCTCCTCTGCTTATCG
TATTTTATCAACAGCACTCTACCATCGATCCTATGTGGAATGTCCGCCACCTTGGTTCCAGT
GCTGGAAAACGATATTCACCTCAGTTTGTAAAGGCTGCCAAGTTACTCCATTGGAATGGACA
TTTGAAGCCATGGGGAAGGACTGCTTCATATACTGATGTTTGGGGAAAAATGGTATATTCCA
GACCCAACAGGCAAATTCAACCTAATCCGAAGATATACCGAGATCTCAACATAAAGTGAAA
CAGAATTTGAACTGTAAGCAAGCATTTCTCAGGAAGTCCTGGAAGATAGCATGCGTGGGAAG
TAACAGTTGCTAGGCTTCAATGCCTATCGGTAGCAAGCCATGGAAAAAGATGTGTCAGCTAG
GTAAAGATGACAACTGCCCTGTCTGGCAGTCAGCTTCCCAGACAGACTATAGACTATAAAT
ATGTCTCCATCTGCCTTACCAAGTGTCTTCTTACTACAATGCTGAATGACTGGAAAGAAGAA
CTGATATGGCTAGTTCAGCTAGCTGGTACAGATAATTCAAACTGCTGTTGGTTTTAATTTT
GTAACCTGTGGCCTGATCTGTAAATAAACTTACATTTTTCAATAGGTAAAAA

FIGURE 108

CTGCAGGTAGACATCTCCACTGCCCAGGAATCACTGAGCGTGCAGACAGCACAGCCTCCTCT
GAAGGCCGGCCATACCAGAGTCCTGCCTCGGCATGGGCCTCACCATTGAGGCAGCTCCACTG
TCTGTGCTGGTCTGAGGGTGCTGCCTGTC**ATG**GGGGCAGCCATCTCCCAGGGGGCCCTCATC
GCCATCGTCTGCAACGGTCTCGTGGGCTTCTTGCTGCTGCTGCTCTGGGTCATCCTCTGCTG
GGCCTGCCATTCTCGTCTGCCGACGTTGACTCTCTCTCTGAATCCAGTCCCAACTCCAGCCC
TGGCCCCCTGTCTGAGAAGGCCCCACCACCCCAGAAGCCCAGCCATGAAGGCAGCTACCTGC
TGCAGCCCTGAAGGCCCTGGCCTAGCCTGGAGCCCAGGACC**TAA**GTCCACCTCACCTAGAG
CCTGGAATTAGGATCCCAGAGTTCAGCCAGCCTGGGGTCCAGAACTCAAGAGTCCGCCTGCT
TGGAGCTGGACCCAGCGGCCCAGAGTCTAGCCAGCTTGGCTCCAATAGGAGCTCAGTGGCCC
TAAGGAGATGGGCCTGGGGTGGGGGCTTATGAGTTGGTGCTAGAGCCAGGGCCATCTGGACT
ATGCTCCATCCCAAGGGCCAAGGGTCAGGGGCCGGGTCCACTCTTCCCTAGGCTGAGCACC
TCTAGGCCCTCTAGGTTGGGGAAGCAAAGTGAACCCATGGCAATAATAGGAGGGTGTCCAG
GCTGGGCCCCCTCCCCTGGTCTCCAGTGTTTGCTGGATAATAAATGGAAGTATGGCTCTAA
AAAAAAAAAAAAAAAAAAAA

109/330

FIGURE 109

MGA AISQGALIAIVCNGLVGFLLLLLWVILCWACHSRLPTLTSLNPVPTPALAPVLRPHH
PRSPAMKAATCCSPEGPWPSLEPRT

109/330

FIGURE 110

GTTTGAATTCCTTCAACTATACCCACAGTCCAAAAGCAGACTCACTGTGTCCCAGGCTACCA
GTTCCCTCCAAGCAAGTCATTTCCCTTATTTAACCGATGTGTCCCTCAAACACCTGAGTGCTA
CTCCCTATTTGCATCTGTTTTGATAAATGATGTTGACACCCTCCACCGAATTCTAAGTGGA
TCATGTCGGGAAGAGATAACAATCCTTGGCCTGTGTATCCTCGCATTAGCCTTGTCTTTGGCC
ATGATGTTTACCTTCAGATTCATCACCACCCTTCTGGTTCACATTTTCATTTTCATTGGTTAT
TTTGGGATTGTTGTTTGTCTGCGGTGTTTTATGGTGGCTGTATTATGACTATACCAACGACC
TCAGCATAGAATTGGACACAGAAAGGGAAAATATGAAGTGCCTGCTGGGGTTTGCTATCGTA
TCCACAGGCATCACGGCAGTGCTGCTCGTCTTGATTTTTGTTCTCAGAAAGAGAATAAAATT
GACAGTTGAGCTTTTCCAAATCACAAATAAAGCCATCAGCAGTGCTCCCTTCCTGCTGTTCC
AGCCACTGTGGACATTTGCCATCCTCATTTTCTTCTGGGTCCTCTGGGTGGCTGTGCTGCTG
AGCCTGGGAAGTGCAGGAGCTGCCAGGTTATGGAAGGCGGCCAAGTGAATATAAGCCCCT
TTCGGGCATTTCGGTACATGTGGTTCGTACCATTTAATTGGCCTCATCTGGACTAGTGAATTCA
TCCTTGCGTGCCAGCAAATGACTATAGCTGGGGCAGTGGTTACTTGTTATTTCAACAGAAGT
AAAAATGATCCTCCTGATCATCCCATCCTTTTCGTCTCTCTCCATTCTCTTCTTCTACCATCA
AGGAACCGTTGTGAAAGGGTCATTTTTTAATCTCTGTGGTGAGGATTCGAGAATCATTGTCA
TGTACATGCAAAACGCACTGAAAGAACAGCAGCATGGTGCATTGTCCAGGTACCTGTTCCGA
TGCTGCTACTGCTGTTTTCTGGTGTCTTGACAAATACCTGCTCCATCTCAACCAGAATGCATA
TACTACAACCTGCTATTAATGGGACAGATTTCTGTACATCAGCAAAAGATGCATTCAAAATCT
TGTTCCAAGAAGTCAAGTCACTTTACATCTATTAACCTGCTTTGGAGACTTCATAATTTTTTCTA
GGAAAGGTGTTAGTGGTGTGTTTCACTGTTTTTTGGAGGACTCATGGCTTTTAACTACAATCG
GGCATTCCAGGTGTGGGCAGTCCCTCTGTTATTGGTAGCTTTTTTTGCCTACTTAGTAGCCC
ATAGTTTTTTTATCTGTGTTTGAAACTGTGCTGGATGCATTTTCCTGTGTTTTGCTGTTGAT
CTGGAAACAAATGATGGATCGTCAGAAAAGCCCTACTTTATGGATCAAGAATTTCTGAGTTT
CGTAAAAAGGAGCAACAAATTAAACAATGCAAGGGCACAGCAGGACAAGCACTCATTAAGGA
ATGAGGAGGGAACAGAACTCCAGGCCATTGTGAGATTAGATACCCATTTAGGTATCTGTACCT
GGAAAACATTTCTTCTAAGAGCCATTTACAGAATAGAAGATGAGACCACTAGAGAAAAGTT
AGTGAATTTTTTTTTTAAAAGACCTAATAAACCTATTCTTCCTCAAAA

111/330

FIGURE 111

MSGRDTILGLCILALALSLAMMFTFRFITLLVHIFISLVILGLLFVCGVLWWLYDYTNDL
SIELDTERENMKCVLGFAIVSTGITAVLLVLI FVLRKRIKLTVELFQITNKAISSAPFLLFQ
PLWTFAILIFFWVLWVAVLLSLGTAGAAQVMEGGQVEYKPLSGIRYMWSYHLIGLIWTSEFI
LACQQMTIAGAVVTCYFNRSKNDPPDHPILSSLSILFFYHQGTVVKGSFLISVVRIPRIIVM
YMQNALKEQQHGALSRYLFRCCYCCFWCLDKYLLHLNQAYTTTAINGTDFCTSAKDAFKIL
SKNSSHFTSINCFGDFIIFLGKVLVVCFTVFGGLMAFNYNRAFQVWAVPLLLVAFFAYLVAH
SFLSVFETVLDALFLCFAVDLETNDGSSEKPYFMDQEFLSFVKRSNKLNNARAQQDKHSLRN
EEGTELQAIVR

099972400
40011226960

FIGURE 117

MQTFTMVLEEIWTSLEFMWFFYALIPCLLTDEVAILPAPQNLSVLSTNMKHLLMWSPVIAPGE
TVYYSVEYQGEYESLYTSHIWIPSSWCSLTEGPECDDITATVPYNLRVRATLGSQTS
SILKHPFNSTILTRPGMEITKDGFLVIELEDLGPQFEFLVAYWRREPGAEHVKMVRSG
GIPVHLETMEPGAAYCVKAQTFVKAIGRYSAFSQTECVQGEAIPVLALFAFVGFMILV
VVPLFVWKMGRLLQYSCCPVVLPDTLKITNSPQKLISCRREEVDACATAVMSPEELLRAWIS

Important features:**Signal peptide:**

amino acids 1-29

Transmembrane domain:

amino acids 230-255

N-glycosylation sites.

amino acids 40-43 and 134-137

Tissue factor proteins homology.

amino acids 92-119

Integrins alpha chain protein homology.

amino acids 232-262

FIGURE 118

TCCTGCTGATGCACATCTGGGTTTGGCAAAAGGAGGTTGCTTCGAGCCGCCCTTTCTAGCTT
CCTGGCCGGCTCTAGAACAATTCAGGCTTCGCTGCGACTAGACCTCAGCTCCAACATATGCA
TTCTGAAGAAAGATGGCTGAGATGACAGAATGCTTTATTTTGGAAAGAAACAATGTTCTAGG
TCAAACCTGAGTCTACCAAATGCAGACTTTTACAATGGTTCTAGAAGAAATCTGGACAAGTCT
TTTCATGTGGTTTTTTCTACGCATTGATTCCATGTTTGCTCACAGATGAAGTGGCCATTCTGC
CTGCCCCCTCAGAACCTCTCTGTACTCTCAACCAACATGAAGCATCTCTTGATGTGGAGCCCA
GTGATCGCGCCTGGAGAAACAGTGTACTATTCTGTCTGAATACCAGGGGGAGTACGAGAGCCT
GTACACGAGCCACATCTGGATCCCCAGCAGCTGGTGCTCACTCACTGAAGGTCCTGAGTGTG
ATGTCACTGATGACATCACGGCCACTGTGCCATACAACCTTTGTGTCTAGGGCCACATTGGGC
TCACAGACCTCAGCCTGGAGCATCCTGAAGCATCCCTTTAATAGAACTCAACCATCCTTAC
CCGACCTGGGATGGAGATCACCAAAGATGGCTTNCACCTGGTTATTGAGCTGGAGGACCTGG
GGCCCCAGTTTGAGTTCCTTGTGGCCTANTGGAGGAGGGGCGAACCCCTTGCGGCGCAAGGG
GTTNGCGAACCCCTTGCGGCCGCTGGGGTATCTCTCGAGAAAAGAGAGGCCCAATATGACCCAC
ATACTCAATATGGACGAANTGCTATTGTCCACCTGTTTGAGTGGCGCTGGGTTGAT

FIGURE 119

CGGACGCGTGGGCCGCCACCTCCGGAACAAGCC**ATG**GTGGCGGCGACGGTGGCAGCGGCGTG
GCTGCTCCTGTGGGCTGCGGCCTGCGCGCAGCAGGAGCAGGACTTCTACGACTTCAAGGCGG
TCAACATCCGGGGCAAACCTGGTGTGCTGGAGAAGTACCGCGGATCGGTGTCCCTGGTGGTG
AATGTGGCCAGCGAGTGCGGCTTCACAGACCAGCACTACCGAGCCCTGCAGCAGCTGCAGCG
AGACCTGGGGCCCCACCACTTTAACGTGCTCGCCTTCCCCTGCAACCAGTTTGGCCAACAGG
AGCCTGACAGCAACAAGGAGATTGAGAGCTTTGCCCCGCCGACCTACAGTGTCTCATTTCCC
ATGTTTAGCAAGATTGCAGTCAACGGTACTGGTGGCCATCCTGCCTTCAAGTACCTGGCCCA
GACTTCTGGGAAGGAGCCCACCTGGAACCTTCTGGAAGTACCTAGTAGCCCCAGATGGAAAGG
TGGTAGGGGCTTGGGACCCAACTGTGTGAGTGGAGGAGGTCAGACCCCAGATCACAGCGCTC
GTGAGGAAGCTCATCCTACTGAAGCGAGAAGACTTA**TAA**ACCACCGCGTCTCCTCCTCCACCA
CCTCATCCCGCCACCTGTGTGGGGCTGACCAATGCAAACCTCAAATGGTGTCTCAAAGGGAG
AGACCCACTGACTCTCCTTCCTTTACTCTTATGCCATTGGTCCCATCATTCTTGTGGGGGAA
AAATTCTAGTATTTTGATTATTTGAATCTTACAGCAACAAATAGGAACTCCTGGCCAATGAG
AGCTCTTGACCAGTGAATCACCAGCCGATACGAACGTCTTGCCAACAAAAATGTGTGGCAAA
TAGAAGTATATCAAGCAATAATCTCCCACCCAAGGCTTCTGTAAACTGGGACCAATGATTAC
CTCATAGGGCTGTTGTGAGGATTAGGATGAAATACCTGTGAAAGTGCCTAGGCAGTGCCAGC
CAAATAGGAGGCATTCAATGAACATTTTTTGCATATAAACCAAAAAATAACTTGTTATCAAT
AAAACTTGCATCCAACATGAATTTCCAGCCGATGATAATCCAGGCCAAAGGTTTAGTTGTT
GTTATTTCTCCTGTATTATTTTCTTCATTACAAAAGAAATGCAAGTTCATTGTAACAATCCA
ACAATACCTCACGATATAAAATAAAAATGAAAGTATCCTCCTCAAAAA

[illegible]

129/330

FIGURE 129

MKIPVLPVAVLLSLLVLHSAQGATLGGPEEESTIENYASRPEAFNTPFLNIDKLRSFAKADE
FLNWHALFESIKRKLPFLNWDAFPKLGKGLRSATPDAQ

bioRxiv preprint doi: <https://doi.org/10.1101/2020.03.10.332660>; this version posted March 10, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

130/330

FIGURE 130

CAGTTCTGAAATCAATGGAGTTAATTTAGGGAATACAAACCAGCCATGGGGGTGGAGATTGC
CTTTGCCTCAGTGATTCTCACCTGCCTCTCCCTTCTGGCAGCAGGAGTCTCCCAGGTTGTTC
TTCTCCAGCCAGTTCCAACCTCAGGAGACAGGTCCCAAGGCCATGGGAGATCTCTCCTGTGGC
TTTGCCGGCCACTCATGAGAGTGTTTTTGTGTAAAGTATTTTTTTAGAATACTGTTGACTTCT
TCATGATTTAATAACCATCCTTTGCGAAGTTTTATGAGGCTTTAGGGGAATGTCAACCCTCA
AATTTTTGTTATACTAGATGGCTTCCATTTACCCACCACTATTTTAAGGTCCCTTTATTTTT
AGGTTCAAGGTTCAATTTGACTTGAGAAAGTGCCCTTCTGCAGCTTCATTGATTTTGTATC
TTCATTAATTGTAACGATTAAAAAGAATAAGAGCACGCAGACCTCTAGGAGAATATTT
TATCCCTGGGTGCCCTGACACATTTATGTAGTGATCCCACAAATGTGATTGTTAATTTAAA
TGTTATTCTAATATTAGTACATTCAGTTGTGATGTAATATGAATAACCAGAATCTATTTCTT
AAAAGTTTGGAGTATATTTTTCAACTAGATATTTGTATAGAAAGACTGAATAGTGATG

bioRxiv preprint doi: <https://doi.org/10.1101/2020.04.15.041111>; this version posted April 15, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

131/330

FIGURE 131

MGVEIAFASVILTCLSLAAGVSQVLLQPVPTQETGPKAMGDLSCGFAGHS

bioRxiv preprint doi: <https://doi.org/10.1101/2021.04.28.440000>; this version posted May 3, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

CACCCCTCCATTTTCTCGCC**ATG**GGCCCCTGCACTGCTCCTGATCCCTGCTGCCCTCGCCTCTT
TCATCCTGGCCTTTGGCACCGGAGTGGAGTTCTGTGCGCTTTACCTCCCTTCGGCCACTTCTT
GGAGGGATCCCGGAGTCTGGTGGTCCGGATGCCCGCCAGGGATGGCTGGCTGCCCTGCAGGA
CCGCAGCATCCTTGCCCCCCTGGCATGGGATCTGGGGCTCCTGCTTCTATTTGTTGGGCAGC
ACAGCCTCATGGCAGCTGAAAGAGTGAAGGCATGGACATCCCGGTACTTTGGGGTCTTTCAG
AGGTCACTGTATGTGGCCTGCACTGCCCTGGCCTTGCACTGGTGATGCGGTACTGGGAGCC
CATACCCAAAGGCCCTGTGTTGTGGGAGGCTCGGGCTGAGCCATGGGCCACCTGGGTGCCGC
TCCTCTGCTTTGTGCTCCATGTCATCTCCTGGCTCCTCATCTTTAGCATCCTTCTCGTCTTT
GACTATGCTGAGCTCATGGGCCTCAAACAGGTATACTACCATGTGCTGGGGCTGGGCGAGCC
TCTGGCCCTGAAGTCTCCCCGGGCTCTCAGACTCTTCTCCCACCTGCGCCACCCAGTGTGTG
TGGAGCTGCTGACAGTGCTGTGGGTGGTGCCTACCCTGGGCACGGACCGTCTCCTCCTTGCT
TTCCTCCTTACCCTCTACCTGGGCCTGGCTCACGGGCTTGATCAGCAAGACCTCCGCTACCT
CCGGGCCCAGCTACAAAGAAAACCTCCACCTGCTCTCTCGGCCCCAGGATGGGGAGGCAGAG**T**
GAGGAGCTCACTCTGGTTACAAGCCCTGTTCTTCTCCTCTCCCACTGAATTCTAAATCCTTAAC
ATCCAGGCCCTGGCTGCTTCATGCCAGAGGCCCAAATCCATGGACTGAAGGAGATGCCCTT
CTACTACTTGAGACTTTATTCTCTGGGTCCAGCTCCATACCCTAAATTCTGAGTTTCAGCCA
CTGAACTCCAAGGTCCACTTCTCACCAGCAAGGAAGAGTGGGGTATGGAAGTCATCTGTCCC
TTCAGTGTGTTAGAGCATGACACTCTCCCCCTCAACAGCCTCCTGAGAAGGAAAGGATCTGCC
CTGACCACTCCCCTGGCACTGTTACTTGCTCTGCGCTCAGGGGTCCCCTTCTGCACCGCT
GGCTTCCACTCCAAGAAGGTGGACCAGGGTCTGCAAGTTCAACGGTCATAGCTGTCCCTCCA
GGCCCCAACCTTGCTCACCCTCCCGGCCCTAGTCTCTGCACCTCCTTAGGCCCTGCCTCT
GGGCTCAGACCCCAACCTAGTCAAGGGGATTCTCCTGCTCTTAACTCGATGACTTGGGGCTC
CCTGCTCTCCCGAGGAAGATGCTCTGCAGGAAAATAAAAGTCAGCCTTTTTTCTAAAAAAA

FIGURE 138

CCTCTGTCCACTGCTTTCGTGAAGACAAGATGAAGTTCACAATTGTCTTTGCTGGACTTCTT
GGAGTCTTTCTAGCTCCTGCCCTAGCTAACTATAATATCAACGTCAATGATGACAACAACAA
TGCTGGAAGTGGGCAGCAGTCAGTGAGTGTCAACAATGAACACAATGTGGCCAATGTTGACA
ATAACAACGGATGGGACTCCTGGAATTCCATCTGGGATTATGGAAATGGCTTTGCTGCAACC
AGACTCTTTCAAAAGAAGACATGCATTGTGCACAAAATGAACAAGGAAGTCATGCCCTCCAT
TCAATCCCTTGATGCACTGGTCAAGGAAAAGAAGCTTCAGGGTAAGGGACCAGGAGGACCAC
CTCCCAAGGGCCTGATGTACTCAGTCAACCCAAACAAAGTCGATGACCTGAGCAAGTTCGGA
AAAAACATTGCAAACATGTGTCTGCTGGGATTCCAACATACATGGCTGAGGAGATGCAAGAGGC
AAGCCTGTTTTTTTTACTCAGGAACGTGCTACACGACCAGTGTACTATGGATTGTGGACATTT
CCTTCTGTGGAGACACGGTGGAGAACTAAACAATTTTTTAAAGCCACTATGGATTTAGTCAT
CTGAATATGCTGTGCAGAAAAAATATGGGCTCCAGTGGTTTTTACCATGTCATTCTGAAATT
TTTCTCTACTAGTTATGTTTGATTTCTTTAAGTTTCAATAAAATCATTTAGCATTGAAAAAAA

FIGURE 139

MKFTIVFAGLLGVFLAPALANYNINVNDDNNNAGSGQQSVSVNNEHNVANVDNNGWDSWNS
IWDYGNNGFAATRLFQKKTCIVHKMNKEVMPSIQSLDALVKEKKLQGKGPGGPPPKGLMYSVN
PNKVDDLKFGKNIANMCRGIPTYMAEEMQEASLFFYSGTCYTTSVLWIVDISFCGDTVEN

Signal Peptide:

amino acids 1-20

N-myristoylation Sites:

amino acids 67-72, 118-123, 163-168

Flavodoxin protein homology:

amino acids 156-174

FIGURE 142

TGGGGATAAAGGAAAAATGGTCAGGTATTAATGGCTTAAAGATTATTGGAAGGGGTTTATCA
TTTTTTGAANNTATTCGGGTCANAATTGNCTTTGAAAAGCATTGCTTTTTACAGAAATATAT
TANCTTTTTAGAGTAATTTCTAGTTTGGATTGTAATATGAAATTATTTAAAAGGGCTTCGCT
CATATATAGGAAAATCGCATATGGTCCTAGTATTAAATTNTTATTGCTTACTGATTTTTTTG
AGTTAAGAGTTGTTATATGNTAGAATATGAGGATGTGAATATAAATAAGAGAAGAAAAAAGA
ATAAAGTAGATTGAGTCTCCAATTTTATGTAAGCTTCAGAAGAAGTGGTTTGTTTACATGCA
AGCTTATAGTTGAAATATTTTTTCAGGAATTACATGAATGACAGTCTTCGAACCAATGTGTTT
GTTCGATTTCAACCAGAGANTATAGCATGTGCTTGCATCTACCTTGCAGNTAGAGCACTTCA
GATTCCGTTGCCAACTNGTCCCCATTGGTTTCTTCTTTTTGGTACTACAGAAGAGGAAATCC
AGGAAATNTGCATAGAAACACTTAGGCTTTATACCAGAAAAAAGCCAACTATGAATTACTG
GAAAAAGAAGTAGAAAAAAGAAAAGTAGCCTTACAAGAAGCCNAATTAAAAGCAAAGGGATT
GAATCCGGATGGAAGTCCAGCCCTTTCAACCCTGGGTGGATTTTCTCC

FIGURE 143

GGCACGAGGCCTCGTGCCAAGCTTGGCACGAGGGTGCACCGCGTTCTCGCACGCGTC**ATGGC**
GGTCCTCGGAGTACAGCTGGTGGTGACCTGCTCACTGCCACCCTCATGCACAGGCTGGCGC
CACACTGCTCCTTCGCGCGCTGGCTGCTCTGTAACGGCAGTTTGTTCGGATACAAGCACCCG
TCTGAGGAGGAGCTTCGGGGCCCTGGCGGGGAAGCCGAGGCCAGAGGCAGGAAAGAGCGGTG
GGCCAATGGCCTTAGTGAGGAGAAGCCACTGTCTGTGCCCCGAGATGCCCCGTTCCAGCTGG
AGACCTGCCCCCTCACGACCGTGATGCCCTGGTCCTGCGCTTCTTCCTGGAGTACCAGTGG
TTTGTGGACTTTGCTGTGTACTCGGGCGGCGTGTACCTCTTCACAGAGGCCTACTACTACAT
GCTGGGACCAGCCAAGGAGACTAACATTGCTGTGTTCTGGTGCCTGCTCACGGTGACCTTCT
CCATCAAGATGTTCTGACAGTGACACGGCTGTACTTCAGCGCCGAGGAGGGGGGTGAGCGC
TCTGTCTGCCTCACCTTTGCCTTCCTCTTCCTGCTGCTGGCCATGCTGGTGCAAGTGGTGCG
GGAGGAGACCCTCGAGCTGGGCCTGGAGCCTGGTCTGGCCAGCATGACCCAGAAGTTAGAGC
CACTTCTGAAGAAGCAGGGCTGGGACTGGGCGCTTCCTGTGGCCAAGCTGGCTATCCGCGTG
GGACTGGCAGTGGTGGGCTCTGTGCTGGGTGCCTTCCTCACCTTCCCAGGCCTGCGGCTGGC
CCAGACCCACCGGGACGCACTGACCATGTGCGAGGACAGACCCATGCTGCAGTTCCTCCTGC
ACACCAGCTTCCTGTCTCCCCTGTTTCATCCTGTGGCTCTGGACAAAGCCCATTGCACGGGAC
TTCCTGCACCAGCCGCCGTTTGGGGAGACGCGTTTCTCCCTGCTGTCCGATTCTGCCTTCGA
CTCTGGGCGCCTCTGGTTGCTGGTGGTGTGCTGTGCCTGCTGCGGCTGGCGGTGACCCGGCCCC
ACCTGCAGGCCTACCTGTGCCTGGCCAAGGCCCGGGTGGAGCAGCTGCGAAGGGAGGCTGGC
CGCATCGAAGCCCGTGAAATCCAGCAGAGGGTGGTCCGAGTCTACTGCTATGTGACCGTGGT
GAGCTTGCAGTACCTGACGCCGCTCATCCTCACCTCAACTGCACACTTCTGCTCAAGACGC
TGGGAGGCTATTCTGGGGCCTGGGCCCAGCTCCTCTACTATCCCCGACCCATCCTCAGCC
AGCGCTGCCCCCATCGGCTCTGGGGAGGACGAAGTCCAGCAGACTGCAGCGCGGATTGCCGG
GGCCCTGGGTGGCCTGCTTACTCCCCTCTTCCTCCGTGGCGTCCTGGCCTACCTCATCTGGT
GGACGGCTGCCTGCCAGCTGCTCGCCAGCCTTTTCGGCCTCTACTTCCACCAGCACTTGGCA
GGCTCC**TAG**CTGCCTGCAGACCCTCCTGGGGCCCTGAGGTCTGTTCTGGGGCAGCGGGACA
CTAGCCTGCCCCCTCTGTTTGCGCCCCCGTGTCCCAGCTGCAAGGTGGGGCCGGACTCCCC
GGCGTTCCCTTCACCACAGTGCTGACCCGCGGCCCCCCTTGGACGCCGAGTTTCTGCCTCA
GAACTGTCTCTCCTGGGCCCAGCAGCATGAGGGTCCCAGAGCCATTGTCTCCGAAGCGTATG
TGCCAGGTTTGAGTGCGGAGGGTGATGCTGGCTGCTCTTCTGAACAAATAAAGGAGCATGCC
GATTTTAA

FIGURE 144

MAVLGVQLVVTLTATLMHRLAPHCSFARWLLCNGSLFRYKHPSEEELRALAGKPRPRGRKE
RWANGLSEEKPLSVPRDAPFQLETCPLTTVDALVLRFFLEYQWFVDFAVYSGGVYLFTEAYY
YMLGPAKETNIAVFWCLLTVTFSIKMFLTVTRLYFSAEEGGERSVCLTFAFLFLLLAMLVQV
VREETLELGLPGLASMTQNLEPLLKKQGWDWALPVAKLAI RVGLAVVGSVLGAFLTFPGLR
LAQTHRDALTMSEDRPMLQFLLHTSFLSPLFILWLWTKPIARDFLHQPPFGETRFSLLSDSA
FDSGRLWLLVVLCLLRLAVTRPHLQAYLCLAKARVEQLRREAGRIEAREIQQRVVRVYCYVT
VVS LQYLTPLIILTLNCTLLLKTLGGYSWGLGPAPLLSPDPSSASAAPIGSGEDEVQQTAAARI
AGALGGLLTPLFLRGVLAYLIWWTAACQLLASLFGLYFHQHLA GS

FIGURE 145

CGTTNGCACGCGTCAATGGCGGTCCTCGGAGTACAGCTGGTGGTGACCCTGCTCACTGCCAC
CCTCATGCACAGGCTGGCGCCACACTGCTCCTTCGCGCGCTGGCTGCTCTGTAAACGGCAGTT
TGTTCCGATACAAGCACCCGTNTTGAGGAGGAGCTTCGGGCCCTGGCGGGGAAGCCGAGGCC
CAGAGGCAGGAAAGAGCGGTGGGCCAATGGCCTTAGTGAGGAGAAGCCACTGTCTGTGCCCC
GAGATGCCCCGTTCCAGCTGGAGACCTGCCCCCTCACGACCGTGGATGCCCTGGTCCTGCGC
TTCTTCCTGGAGTACCAGTGGTTTGTGGACTTTGCTGTGTACTCGGGCGGCGTGTACCTCTT
CACAGAGGCCTACTACTACATGCTGGGACCAGCCAAGGAGACTAACATTGCTGTGTTCTGGT
GCCTGCTCACAGTGACCTTCTCCATCAAGATGTTCTGACAGTGACACGGCTGTACTTCAGC
GCCGAGGAGGGGGGTGAGCGCTCTGTCTGCCTCACCTTTGCCTTCCTCTTCCTGCTGCTGGC
CATGCTGGTGCAAGCG

FIGURE 148

CCAAAGTGATCATTTGAAAAAGAGATATCCACATCTTCAAGCCCATATAAAGGATAGAAGCT
GCACAGGGCAGCTTTACTTACTCCAGCACCTTCCTCTCCCAGGCAA**ATG**GTGCTGACCATCT
TTGGGATACAATCTCATGGATACGAGGTTTTTAACATCATCAGCCCAAGCAACAATGGTGGC
AATGTTTCAGGAGACAGTGACAATTGATAATGAAAAAATACCGCCATCGTTAACATCCATGC
AGGATCATGCTCTTCTACCACAATTTTTGACTATAAACATGGCTACATTGCATCCAGGGTGC
TCTCCCGAAGAGCCTGCTTTATCCTGAAGATGGACCATCAGAACATCCCTCCTCTGAACAAT
CTCCAATGGTACATCTATGAGAAACAGGCTCTGGACAACATGTTCTCCAACAAATACACCTG
GGTCAAGTACAACCCTCTGGAGTCTCTGATCAAAGACGTGGATTGGTTCCTGCTTGGGTCAC
CCATTGAGAACTCTGCAAACATATCCCTTTGTATAAGGGGGGAAGTGGTTGAAAACACACAT
AATGTCGGTGCTGGAGGCTGTGCAAAGGCTGGGCTCCTGGGCATCTTGGGAATTTCAATCTG
TGCAGACATTCATGTT**TAG**GATGATTAGCCCTCTTGTTTTATCTTTTCAAAGAAATACATCC
TTGGTTTACACTCAAAAGTCAAATTAAATTCTTTCCCAATGCCCCAACTAATTTTGAGATTC
AGTCAGAAAATATAAATGCTGTATTTATA

149/330

FIGURE 149

MKILVAFLVVLTIFFGIQSHGYEVFNIIISPSNNGGNVQETVTIDNEKNTAIVNIHAGSCSSTT
IFDYKHGYIASRVLSRRACFILKMDHQNIPLNNLQWYIYEKQALDNMFSNKYTWVKYNPLE
SLIKDVDWFLLGSPIEKLCKHIPLYKGEVVENTHNVGAGGCAKAGLLGILGISICADIHV

149/330

150/330

FIGURE 150

GGCACGAGCCAGGAACTAGGAGGTTCTCACTGCCCCGAGCAGAGGCCCTACACCCACCGAGGC
ATGGGGCTCCCTGGGCTGTTCTGCTTGGCCGTGCTGGCTGCCAGCAGCTTCTCCAAGGCACG
GGAGGAAGAAATTACCCCTGTGGTCTCCATTGCCTACAAAGTCCTGGAAGTTTTCCCCAAAG
GCCGCTGGGTGCTCATAACCTGCTGTGCACCCCAGCCACCACCGCCCATCACCTATTCCCTC
TGTGGAACCAAGAACATCAAGGTGGCCAAGAAGGTGGTGAAGACCCACGAGCCGGCCTCCTT
CAACCTCAACGTCACACTCAAGTCCAGTCCAGACCTGCTCACCTACTTCTGCCGGGCGTCCT
CCACCTCAGGTGCCCATGTGGACAGTGCCAGGCTACAGATGCACTGGGAGCTGTGGTCCAAG
CCAGTGTCTGAGCTGCGGGCCAACCTTCACTCTGCAGGACAGAGGGGCAGGCCCCAGGGTGGA
GATGATCTGCCAGGCGTCCTCGGGCAGCCCACCTATCACCAACAGCCTGATCGGGAAGGATG
GGCAGGTCCACCTGCAGCAGAGACCATGCCACAGGCAGCCTGCCAACTTCTCCTTCCTGCCG
AGCCAGACATCGGACTGGTTCTGGTGCCAGGCTGCAAACAACGCCAATGTCCAGCACAGCGC
CCTCACAGTGGTGCCCCCAGGTGGTGACCAGAAGATGGAGGACTGGCAGGGTCCCCTGGAGA
GCCCCATCCTTGCCTTGCCGCTCTACAGGAGCACCCGCCGTCTGAGTGAAGAGGAGTTTGGG
GGGTTTCAGGATAGGGAATGGGGAGGTCAGAGGACGCAAAGCAGCAGCCATG**TAG**AATGAACC
GTCCAGAGAGCCAAGCACGGCAGAGGACTGCAGGCCATCAGCGTGCACTGTTTCGTATTTGGA
GTTTCATGCAAAATGAGTGTGTTTTAGCTGCTCTTGCCACAAAAAAAAAAAAAAAAAAAAAA

150/330

FIGURE 153

MAAAAATKILLCLPLLLLLSGWSRAGRADPHSLCYDITVIPKFRPGPRWCAVQGQVDEKTF
HYDCGNKTVTPVSPPLGKKLNVTAWKAQNPVLREVVDILTEQLRDIQLENYTPKEPLTLQAR
MSCEQKAEGHSSGSWQFSFDGQIFLLFDSEKRMWTTVHPGARKMKEKWENDKVVAMSFHYFS
MGDCIGWLEDFLMGMDSTLEPSAGAPLAMSSGTTQLRATATTLILCCLLIILPCFILPGI

Important features:**Signal peptide:**

amino acids 1-25

Transmembrane domain:

amino acids 224-246

N-glycosylation site.

amino acids 68-72, 82-86

N-myristoylation site.

amino acids 200-206, 210-216

Amidation site.

amino acids 77-81

FIGURE 154

GGGAAAGCCATTTGCGAAAACCCATCTATACAACTATATATTTTCATTTCTGCTGCTAGCTG
CCTTGGGCCTCACAAATTTTCATTCTGTTTTCTGACTTTCAAGTTATATACCGTGGA**ATG**GAG
TTGATCCCAACCATAACATCGTGGAGGGTTTTAATTTTGGTGGTAGCCCTCACCCAATTCTG
GTGTGGCTTTCTTTGCAGAGGATTCCACCTTCAAATCATGAACTCTGGCTGTTGATCAAAA
GAGAATTTGGATTCTACTCTAAAAGTCAATATAGGACTTGGCAAAAGAAGCTAGCAGAAGAC
TCAACCTGGCCTCCCATAAACAGGACAGATTATTCAGGTGATGGCAAAAATGGATTCTACAT
CAACGGAGGCTATGAAAGCCATGAACAGATTCCAAAAAGAAAACCTCAAATTGGGAGGCCAAC
CCACAGAACAGCATTCTGCGGCCAGGCTG**TAA**TCAGAATTGTCGTCGTACATGCTCAACAGC
ATTGCTTTTTTCCCCAAAATTAACACATTGTGGAGAAGTGATGATACTCTCCCCTTACCTTT
CCTCTCTCCATTCAAGCATTCAAAGTATATTTTCAATGAATTAAACCTTGCAGCAAGGGACC
TTAGATAGGCTTATTCTGACTGTATGCTTTACCAATGAGAGAAAAAAATGCATTTTCTGTAT
CATCCTTTTCAATAAACTGTATTTCATTTTGAAAAAAAAAAAAAAAAAAAAA

155/330

FIGURE 155

MELIPTITSWRVLILVVALTQFWCGFLCRGFHLQNHFWLLIKREFGFYSKSQYRTWQKKLA
EDSTWPPINRTDYSGDGKNGFYINGGYESHEQIPKRKLKLGGQPTQHFWARL

bioRxiv preprint doi: <https://doi.org/10.1101/2021.04.28.440000>; this version posted April 28, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 156

GTTCTCCTTTCCGAGCCAAAATCCCAGGCGATGGTGAATTATGAACGTGCCACACC**ATGAAG**
 CTCTTGTGGCAGGTAACTGTGCACCACCACACCTGGAATGCCATCCTGCTCCCGTTCTGTCTA
 CCTCACGGCGCAAGTGTGGATTCTGTGTGCAGCCATCGCTGCTGCCGCTCAGCCGGGCCCC
 AGAACTGCCCCCTCCGTTTGCTCGTGCAGTAACCAAGTTCAGCAAGGTGGTGTGCACGCGCCGG
 GGCCTCTCCGAGGTCCCGCAGGGTATTCCCTCGAACACCCGGTACCTCAACCTCATGGAGAA
 CAACATCCAGATGATCCAGGCCGACACCTTCCGCCACCTCCACCACCTGGAGGTCTTGCAGT
 TGGGCAGGAACTCCATCCGGCAGATTGAGGTGGGGGCCTTCAACGGCCTGGCCAGCCTCAAC
 ACCCTGGAGCTGTTTCGACAACTGGCTGACAGTCATCCCTAGCGGGGCCTTTGAATACCTGTC
 CAAGCTGCGGGAGCTCTGGCTTCGCAACAACCCCATCGAAAGCATCCCCTCTTACGCCTTCA
 ACCGGGTGCCCTCCCTCATGCGCTGGACTTGGGGGAGCTCAAGAAGCTGGAGTATATCTCT
 GAGGGAGCTTTTGAGGGGCTGTTCAACCTCAAGTATCTGAACTTGGGCATGTGCAACATTAA
 AGACATGCCCAATCTCACCCCCCTGGTGGGGCTGGAGGAGCTGGAGATGTCAGGGAACCACT
 TCCCTGAGATCAGGCCTGGCTCCTTCCATGGCCTGAGCTCCCTCAAGAAGCTCTGGGTCTATG
 AACTCACAGGTGAGCCTGATTGAGCGGAATGCTTTTGACGGGCTGGCTTCACTTGTGGAAC
 CAACTTGGCCCAACAATAACCTCTCTTCTTTGCCCCATGACCTCTTTACCCCGCTGAGGTACC
 TGGTGGAGTTGCATCTACACCACAACCTTGGAACTGTGATTGTGACATTCTGTGGCTAGCC
 TGGTGGCTTCGAGAGTATATACCCACCAATTCCACCTGCTGTGGCCGCTGTCTATGCTCCCAT
 GCACATGCGAGGCCGCTACCTCGTGGAGGTGGACCAGGCCTCCTTCCAGTGCTCTGCCCCCT
 TCATCATGGACGCACCTCGAGACCTCAACATTTCTGAGGGTCGGATGGCAGAACTTAAGTGT
 CGGACTCCCCCTATGTCTCCGTGAAGTGGTTGCTGCCCAATGGGACAGTGCTCAGCCACGC
 CTCCCGCCACCCAAGGATCTCTGTCTCAACGACGGCACCTTGAACTTTTCCACGTGCTGC
 TTTCAGACACTGGGGTGTACACATGCATGGTGACCAATGTTGCAGGCAACTCCAACGCCTCG
 GCCTACCTCAATGTGAGCACGGCTGAGCTTAACACCTCCAACCTACAGCTTCTTACCACAGT
 AACAGTGGAGACCACGGAGATCTCGCCTGAGGACACAACGCGAAAGTACAAGCCTGTTCCCTA
 CCACGTCCACTGGTTACCAGCCGGCATATACCACCTCTACCACGGTGCTCATTACAGACTACC
 CGTGTGCCCAAGCAGGTGGCAGTACCCGCGACAGACACCACTGACAAGATGCAGACCAGCCT
 GGATGAAGTCATGAAGACCACCAAGATCATCATTGGCTGCTTTGTGGCAGTGACTCTGCTAG
 CTGCCGCCATGTTGATTGTCTTCTATAAACTTCGTAAGCGGCACCAGCAGCGGAGTACAGTC
 ACAGCCGCCCGGACTGTTGAGATAATCCAGGTGGACGAAGACATCCAGCAGCAACATCCGC
 AGCAGCAACAGCAGCTCCGTCCGGTGTATCAGGTGAGGGGGCAGTAGTGCTGCCACAATTC
 ATGACCATATTAACATAACACCTACAAACCAGCACATGGGGCCCACTGGACAGAAAACAGC
 CTGGGGAACTCTCTGCACCCACAGTCACCACTATCTCTGAACCTTATATAATTACAGACCA
 TACCAAGGACAAGGTACAGGAACTCAAATA**TGA**CTCCCCTCCCCCAAAAACCTTATAAAAT
 GCAATAGAATGCACACAAAGACAGCAACTTTTGTACAGAGTGGGGAGAGACTTTTTCTTGTA
 TATGCTTATATATTAAGTCTATGGGCTGGTTAAAAAAAACAGATTATATTAATTTAAAGA
 CAAAAAGTCAAAACA

FIGURE 163

MAQAVWSRLGRILWLACLLPWAPAGVAAGLYELNLTTDSPATTGAVVTISASLVAKDNGSLA
LPADAHLYRFHWIHTPLVLTGKMEKGLSSTIRVVGHVPGEFPVSVWVTAADCWMCQPVARGF
VVLPITEFLVGDLVVTQNTSLPWSSYLTKTVLKVSFLLHDPSNFLKTALFLYSWDFGDGTQ
MVTEDSVVYYNYSIIGTFTVKLKVVAEWEEVEPDATRAVKQKTGDFSASLKLQETLRGIQVL
GPTLIQTFQKMTVTLNFLGSPPLTVCWRLKPECLPLEEGECHPVSVASTAYNLTHTFRDPGD
YCFSSIRAENIISKTHQYHKIQVWPSRIQPAVFAFPCATLITVMLAFIMYMTLRNATQQKDMV
ENPEPPSGVRCCCQCCGPFLLLETPSEYLEIVRENHGLLPPLYKSVKTYTV

Important features of the protein:**Signal peptide:**

amino acids 1-24

Transmembrane domain:

amino acids 339-362

N-glycosylation sites.

amino acids 34-37, 58-61, 142-145, 197-200, 300-303 and 364-367

FIGURE 168

GCCCCTCGAAACCAGGACTCCAGCACCTCTGGTCCCGCCCTCACCCGGACCCCTGGCCCTCA
CGTCTCCTCCAGGGATGGCGCTGGCGGCTTTGATGATCGCCCTCGGCAGCCTCGGCCTCCAC
ACCTGGCAGGCCCAGGCTGTTCCCACCATCCTGCCCCCTGGGCCTGGCTCCAGACACCTTTGA
CGATACCTATGTGGGTTGTGCAGAGGAGATGGAGGAGAAGGCAGCCCCCTGCTAAAGGAGG
AAATGGCCCACCATGCCCTGCTGCGGGAATCCTGGGAGGCAGCCCAGGAGACCTGGGAGGAC
AAGCGTCGAGGGCTTACCTTGCCCCCTGGCTTCAAAGCCCAGAATGGAATAGCCATTATGGT
CTACACCAACTCATCGAACACCTTGTACTGGGAGTTGAATCAGGCCGTGCGGACGGGCGGAG
GCTCCCGGGAGCTCTACATGAGGCACTTTCCCTTCAAGGCCCTGCATTTCTACCTGATCCGG
GCCCTGCAGCTGCTGCGAGGCAGTGGGGGCTGCAGCAGGGGACCTGGGGAGGTGGTGTTCGG
AGGTGTGGGCAGCCTTCGCTTTGAACCCAAGAGGCTGGGGGACTCTGTCCGCTTGGGCCAGT
TTGCCTCCAGCTCCCTGGATAAGGCAGTGGCCACAGATTTGGGGAGAAGAGGCGGGGCTGT
GTGTCTGCGCCAGGGGTGCAGCTAGGGTCACAATCTGAGGGGGCCTCCTCTCTGCCCCCTG
GAAGACTCTGCTCTTGGCCCCCTGGAGAGTTCCAGCTCTCAGGGGTTGGGCCCTGAAAGTCCA
ACATCTGCCACTTAGGAGCCCTGGGAACGGGTGACCTTCATATGACGAAGAGGCACCTCCAG
CAGCCTTGAGAAGCAAGAACATGGTTCCGGACCCAGCCCTAGCAGCCTTCTCCCCAACCAGG
ATGTTGGCCTGGGGAGGCCACAGCAGGGCTGAGGGAACTCTGCTATGTGATGGGGACTTCCT
GGGACAAGCAAGGAAAGTACTGAGGCAGCCACTTGATTGAACGGTGTGCAATGTGGAGACA
TGGAGTTTTATTGAGGTAGCTACGTGATTAAATGGTATTGCAGTGTGGA

FIGURE 172

CTGGTTCCCCAACATGCCTCACCCCTCATCTATATCCTTTGGCAGCTCACAGGGTCAGCAGCC
TCTGGACCCCGTGAAAGAGCTGGTCGGTTCCGTTGGTGGGGCCGTGACTTTCCCCCTGAAGTC
CAAAGTAAAGCAAGTTGACTCTATTGTCTGGACCTTCAACACAACCCCTCTTGTCAACATAC
AGCCAGAAGGGGGCCTATCATAGTGACCCAAAATCGTAATAGGGAGAGAGTAGACTTCCCA
GATGGAGGCTACTCCCTGAAGCTCAGCAAACCTGAAGAAGAATGACTCAGGGATCTACTATGT
GGGGATATACAGCTCATCACTCCAGCAGCCCTCCACCCAGGAGTACGTGCTGCATGTCTACG
AGCACCTGTCAAAGCCTAAAGTCACCATGGGTCTGCAGAGCAATAAGAATGGCACCTGTGTG
ACCAATCTGACATGCTGCATGGAACATGGGGAAGAGGATGTGATTTATACCTGGAAGGCCCT
GGGGCAAGCAGCCAATGAGTCCCATAATGGGTCCATCCTCCCCATCTCCTGGAGATGGGGAG
AAAGTGATATGACCTTCATCTGCGTTGCCAGGAACCCTGTCAGCAGAACTTCTCAAGCCCC
ATCCTTGCCAGGAAGCTCTGTGAAGGTGCTGCTGATGACCCAGATTCTCCATGGTCCTCCT
GTGTCTCCTGTTGGTGCCCCCTCCTGCTCAGTCTCTTTGTACTGGGGCTATTTCTTTGGTTTC
TGAAGAGAGAGAGACAAGAAGAGTACATTGAAGAGAAGAAGAGAGTGGACATTTGTCTGGGAA
ACTCCTAACATATGCCCCCATCTGAGAGAGAACACAGAGTACGACACAATCCCTCACACTAA
TAGAACAATCCTAAAGGAAGATCCAGCAAATACGGTTTACTCCACTGTGGAAATACCGAAAA
AGATGGAAAATCCCCACTCACTGCTCACGATGCCAGACACACCAAGGCTATTTGCCTATGAG
AATGTTATCTAGACAGCAGTGCACTCCCCTAAGTCTCTGCTCAAAAAAAAAAAAAAAAAAAAA

FIGURE 173

GAAAGACGTGGTCCTGACAGACAGACAATCCTATTCCCTACCAAAATGAAGATGCTGCTGCT
GCTGTGTTTGGGACTGACCCTAGTCTGTGTCCATGCAGAAGAAGCTAGTTCTACGGGAAGGA
ACTTTAATGTAGAAAAGATTAATGGGGAATGGCATACTATTATCCTGGCCTCTGACAAAAGA
GAAAAGATAGAAGAACATGGCAACTTTAGACTTTTTCTGGAGCAAATCCATGTCTTGGAGAA
TTCCTTAGTTCTTAAAGTCCATACTGTAAGAGATGAAGAGTGCTCCGAATTATCTATGGTTG
CTGACAAAACAGAAAAGGCTGGTGAATATTCTGTGACGTATGATGGATTCAATACATTTACT
ATACCTAAGACAGACTATGATAACTTTCTTATGGCTCACCTCATTAACGAAAAGGATGGGGA
AACCTTCCAGCTGATGGGGCTCTATGGCCGAGAACCAGATTTGAGTTCAGACATCAAGGAAA
GGTTTGCACAACTATGTGAGGAGCATGGAATCCTTAGAGAAAATATCATTGACCTATCCAAT
GCCAATCGCTGCCTCCAGGCCCGAGAATGAAGAATGGCCTGAGCCTCCAGTGTTGAGTGGAC
ACTTCTCACCAGGACTCCACCATCATCCCTTCCTATCCATACAGCATCCCCAGTATAAATTC
TGTGATCTGCATTCCATCCTGTCTCACTGAGAAGTCCAATTCCAGTCTATCAACATGTTACC
TAGGATACCTCATCAAGAATCAAAGACTTCTTTAAATTTCTCTTTGATACACCCTTGACAAT
TTTTCATGAAATTATTCCTCTTCCTGTTCAATAAATGATTACCCTTGCACTTAA

176/330

FIGURE 176

MTCCEGWTSCNGFSLLVLLLLGVVLNAIPLIVSLVEEDQFSQNPISCFEWWFPGIIGAGLMA
IPATTMSLTARKRACCNNRTGMFLSSFFSVITVIGALYCMLISIQALLKGPLMCNSPSNSNA
NCEFSLKNISDIHPESFNLQWFFNDSCAPPTGFNKPTSNDTMASGWRASSFHFDSEENKHRL
IHFSVFLGLLLVGILEVLFGLSQIVIGFLGCLCGVSKRRSQIV

4064.4 = 426360

178/330

FIGURE 178

MRLSVCLLMVSLALCCYQAHALVCPAVASEITVFLFLSDAAVNLQVAKLNPPPEALAAKLEV
KHCTDQISFKKRLSLKKSWWK

095927 22 263550

FIGURE 179

ATCCGTTCTCTGCGCTGCCAGCTCAGGTGAGCCCTCGCCAAGGTGACCTCGCAGGACACTGG
TGAAGGAGCAGTGAGGAACCTGCAGAGTCACACAGTTGCTGACCAATTGAGCTGTGAGCCTG
GAGCAGATCCGTGGGCTGCAGACCCCCGCCCCAGTGCCTCTCCCCCTGCAGCCCTGCCCCCTC
GAACTGTGACATGGAGAGAGTGACCCTGGCCCTTCTCCTACTGGCAGGCCTGACTGCCTTGG
AAGCCAATGACCCATTTGCCAATAAAGACGATCCCTTCTACTATGACTGGAAAAACCTGCAG
CTGAGCGGACTGATCTGCGGAGGGCTCCTGGCCATTGCTGGGATCGCGGCAGTTCTGAGTGG
CAAATGCAAATACAAGAGCAGCCAGAAGCAGCACAGTCCTGTACCTGAGAAGGCCATCCCAC
TCATCACTCCAGGCTCTGCCACTACTTGCTTGAGCACAGGACTGGCCTCCAGGGATGGCCTGA
AGCCTAACACTGGCCCCCAGCACCTCCTCCCCTGGGAGGCCTTATCCTCAAGGAAGGACTTC
TCTCCAAGGGCAGGCTGTTAGGCCCCCTTCTGATCAGGAGGCTTCTTTATGAATTAACTCG
CCCCACCACCCCCTCA

FIGURE 181

GGAGAAGAGGTTGTGTGGGACAAGCTGCTCCCGACAGAAGG**ATG**TCGCTGCTGAGCCTGCCC
 TGGCTGGGCCTCAGACCGGTGGCAATGTCCCCATGGCTACTCCTGCTGCTGGTTGTGGGCTC
 CTGGCTACTCGCCCGCATCCTGGCTTGGACCTATGCCTTCTATAACAACCTGCCGCCGGCTCC
 AGTGTTTCCACAGCCCCCAAACGGAACTGGTTTTTGGGGTCACCTGGGCCTGATCACTCCT
 ACAGAGGAGGGCTTGAAGGACTCGACCCAGATGTCGGCCACCTATTCCCAGGGCTTTACGGT
 ATGGCTGGGTCCCATCATCCCCTTCATCGTTTTATGCCACCCTGACACCATCCGGTCTATCA
 CCAATGCCTCAGCTGCCATTGCACCCAAGGATAATCTCTTCATCAGGTTCTGAAGCCCTGG
 CTGGGAGAAGGGATACTGCTGAGTGGCGGTGACAAGTGGAGCCGCCACCGTCGGATGCTGAC
 GCCCCGCTTCCATTTCAACATCCTGAAGTCCTATATAACGATCTTCAACAAGAGTGCAAACA
 TCATGCTTGACAAGTGGCAGCACCTGGCCTCAGAGGGCAGCAGTCGTCTGGACATGTTTGAG
 CACATCAGCCTCATGACCTTGGACAGTCTACAGAAATGCATCTTCAGCTTTGACAGCCATTG
 TCAGGAGAGGCCCAGTGAATATATTGCCACCATCTTGGAGCTCAGTGCCCTTGTAGAGAAAA
 GAAGCCAGCATATCCTCCAGCACATGGACTTTCTGTATTACCTCTCCCATGACGGGCGGCGC
 TTCCACAGGGCCTGCCGCTGGTGCATGACTTCACAGACGCTGTCATCCGGGAGCGGCGTCCG
 CACCCTCCCCACTCAGGGTATTGATGATTTTTTCAAAGACAAAGCCAAGTCCAAGACTTTGG
 ATTTTCATTGATGTGCTTCTGCTGAGCAAGGATGAAGATGGGAAGGCATTGTCAGATGAGGAT
 ATAAGAGCAGAGGCTGACACCTTCATGTTTGGAGGCCATGACACCACGGCCAGTGGCCTCTC
 CTGGGTCTGTACAACCTTGCAGGGCACCCAGAATACCAGGAGCGCTGCCGACAGGAGGTGC
 AAGAGCTTCTGAAGGACCGCGATCCTAAAGAGATTGAATGGGACGACCTGGCCCAGCTGCCC
 TTCCTGACCATGTGCGTGAAGGAGAGCCTGAGGTTACATCCCCCAGCTCCCTTCATCTCCCG
 ATGCTGCACCCAGGACATTGTTCTCCAGATGGCCGAGTCATCCCCAAAGGCATTACCTGCC
 TCATCGATATTATAGGGGTCCATCACAACCCAACTGTGTGGCCGGATCCTGAGGTCTACGAC
 CCCTTCCGCTTTGACCCAGAGAACAGCAAGGGGAGGTCACCTCTGGCTTTTATTCCCTTTCTC
 CGCAGGGCCCAGGAAGTGCATCGGGCAGGCGTTCCGCATGGCGGAGATGAAAGTGGTCCTGG
 CGTTGATGCTGCTGCACTTCCGGTTCTGCCAGACCACACTGAGCCCCGAGGAAGCTGGAA
 TTGATCATGCGCGCCGAGGGCGGGCTTTGGCTGCGGGTGGAGCCCCTGAATGTAGGCTTGCA
 G**TGA**CTTTCTGACCCATCCACCTGTTTTTTTGCAGATTGTCATGAATAAAACGGTGCTGTCAAA

FIGURE 183

CAACAGAAGCCAAGAAGGAAGCCGCTCTATCTTGTGGCGATC**ATG**TATAAGCTGGCCTCCTGC
TGTTTGCTTTTTCACAGGATTCTTAAATCCTCTCTTATCTCTTCCTCTCCTTGACTCCAGGGA
AATATCCTTTCAACTCTCAGCACCTCATGAAGACGCGCGCTTAACTCCGGAGGAGCTAGAAA
GAGCTTCCCTTCTACAGATATTGCCAGAGATGCTGGGTGCAGAAAGAGGGGATATTCTCAGG
AAAGCAGACTCAAGTACCAACATTTTTTAACCCAAGAGGAAATTTGAGAAAGTTTCAGGATTT
CTCTGGACAAGATCCTAACATTTTACTGAGTCATCTTTTGGCCAGAATCTGGAAACCATAACA
AGAAACGTGAGACTCCTGATTGCTTCTGGAAATACTGTGT**TGA**AGTGAAATAAGCATCTGT
TAGTCAGCTCAGAAACACCCATCTTAGAATATGAAAAATAACACAATGCTTGATTTGAAAAC
AGTGTGGAGAAAACTAGGCAAACTACACCCTGTTTCATTGTTACCTGGAAAATAAATCCTCT
ATGTTTTGCACAAAAAAAAAAAAAAAAA

185/330

FIGURE 185

GAACATTTT TAGTTCCCAAGGAATGTACATCAGCCCCACGGAAGCTAGGCCACCTCTGGGAT
GGGGTTGCTGGTTTAAACAAACGCCAGTCATCCTATATAAGGACCTGACAGCCACCAGGCA
CCACCTCCGCCAGGAACTGCAGGCCCCACCTGTCTGCAACCCAGCTGAGGCCATGCCCTCCCC
AGGGACCGTCTGCAGCCTCCTGCTCCTCGGCATGCTCTGGCTGGACTTGGCCATGGCAGGCT
CCAGCTTCCTGAGCCCTGAACACCAGAGAGTCCAGCAGAGAAAGGAGTCGAAGAAGCCACCA
GCCAAGCTGCAGCCCCGAGCTCTAGCAGGCTGGCTCCGCCCCGGAAGATGGAGGTCAAGCAGA
AGGGGCAGAGGATGAACTGGAAGTCCGGTTCAACGCCCCCTTTGATGTTGGAATCAAGCTGT
CAGGGGTT CAGTACCAGCAGCACAGCCAGGCCCTGGGGAAGTTTCTTCAGGACATCCTCTGG
GAAGAGGCCAAAGAGGGCCCCAGCCGACAAGTGATCGCCACAAGCCTTACTCACCTCTCTCT
AAGTTTAGAAGCGCTCATCTGGCTTTTCGCTTGCTTCTGCAGCAACTCCCACGACTGTTGTA
CAAGCTCAGGAGGCGAATAAATGTTCAAACCTGTA

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

188/330

FIGURE 188

MNTWLLFLPLFPVQVQTLIVVIIGMLVLLLDLGLVHLGQLLIFHIYLSMSPTLSPRSPQGW
VVRAAHLTPLLEYVPNPEPPTPGARVFVPRVRMCSGSASPRSEIMDKKGKSQEEIKSMRTQQ
AQQEAELTPRPAGVVPGA

bioRxiv preprint doi: <https://doi.org/10.1101/2021.04.28.440000>; this version posted May 3, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 189

GGAGTGCAGATGGCATCCTTCGGTTCTTCCAGACAAGCTGCAAGACGCTGACCATGGCCAAG
ATGGAGCTCTCGAAGGCCTTCTCTGGCCAGCGGACACTCCTATCTGCCATCCTCAGCATGCT
ATCACTCAGCTTCTCCACAACATCCCTGCTCAGCAACTACTGGTTTGTGGGCACACAGAAGG
TGCCCAAGCCCCTGTGCGAGAAAGGTCTGGCAGCCAAGTGCTTTGACATGCCAGTGTCCCTG
GATGGAGATACCAACACATCCACCCAGGAGGTGGTACAATACTGGGAGACTGGGGATGA
CCGGTTCTCCTTCCGGAGCTTCCGGAGTGGCATGTGGCTATCCTGTGAGGAACTGTGGAAG
AACCAGGGGAGAGGTGCCGAAGTTTCATTGAACCTACACCACCAGCCAAGAGAGGTGAGAAA
GGACTACTGGAATTTGCCACGTTGCAAGGCCCATGTCACCCCACTCTCCGATTTGGAGGGAA
GCGGTTGATGGAGAAGGCTTCCCTCCCTCCCTCCCTTGGGGCTTTGTGGCAAAAATCCTA
TGTTTATCCCTGGGAACGCAGATCACCTACATCGGACTTCAATTCATCAGCTTCCTCCTGCT
ACTAACAGACTTGCTACTCACTGGGAACCCTGCCTGTGGGCTCAAAGTGGAGCGCCTTTGCTG
CTGTTTCCTCTGTCCTGTCAGGTCTCCTGGGGATGGTGGCCACATGATGTATTACAAAGTC
TTCCAAGCGACTGTCAACTTGGGTCCAGAAGACTGGAGACCACATGTTTGGAAATTATGGCTG
GGCCTTCTACATGGCCTGGCTCTCCTTCACCTGCTGCATGGCGTCGGCTGTCACCACCTTCA
ACACGTACACCAGGATGGTGCTGGAGTTCAAGTGCAAGCATAGTAAGAGCTTCAAGGAAAAC
CCGAAGTGCCTACCACATCACCATCAGTGTTTCCCTCGGCGGCTGTCAAGTGCAGCCCCCAC
CGTGGGTCCCTTTGACCAGCTACCACCAGTATCATAATCAGCCCATCCACTCTGTCTCTGAGG
GAGTCGACTTCTACTCCGAGCTGCGGAACAAGGGATTTCAAAGAGGGGCCAGCCAGGAGCTG
AAAGAAGCAGTTAGGTCATCTGTAGAGGAAGAGCAGTGTTAGGAGTTAAGCGGGTTTGGGGA
GTAGGCTTGAGCCCTACCTTACACGTCTGCTGATTATCAACATGTGCTTAAGCCAACATCCG
TCTCTTGAGCATGGTTTTTTAGAGGCTACGAATAAGGCTATGAATAAGGGTTATCTTTAAGTC
CTAAGGGATTCTGGGTGCCACTGCTCTCTTTTCTCTACAGCTCCATCTTGTTTCACCCAC
CCCACATCTCACACATCCAGAATTCCCTTCTTTACTGATAGTTTCTGTGCCAGGTTCTGGGC
TAAACCATGGAGATAAAAAGAAGAGTAAAATACACTTCCCGACCTTAAGGATCTGAAA

FIGURE 193

GTAGCGCGTCTTGGGTCTCCCGGCTGCCGCTGCTGCCGCCGCCGCTCGGGTCGTGGAGCCAGGAGCGACGTCA
 CCGCC**ATGG**CAGGCATCAAAGCTTTGATTAGTTTGTCTTTGGAGGAGCAATCGGACTGATGTTTTTGATGCTT
 GGATGTGCCCTTCCAATATACAACAAATACTGGCCCTCTTTGTTCTATTTTTTACATCCTTTCACCTATTCC
 ATACTGCATAGCAAGAAGATTAGTGGATGATACAGATGCTATGAGTAACGCTTGTAAGGAACCTTGCCATCTTTC
 TTACAACGGGCATTGTCGTGTGAGCTTTTGGACTCCCTATTGTATTTGCCAGAGCACATCTGATTGAGTGGGA
 GCTTGTGCACTTGTCTCACAGGAAACACAGTCATCTTTGCAACTATACTAGGCTTTTTCTTGGTCTTTGGAAG
 CAATGACGACTTCAGCTGGCAGCAGTGG**TGA**AAAGAAATTACTGAACTATTGTCAAATGGACTTCCTGTCAATT
 GTTGGCCATTACGCACACAGGAGATGGGGCAGTTAATGCTGAATGGTATAGCAAGCCTCTTGGGGGTATTTTA
 GGTGCTCCCTTCTCACTTTTATTGTAAGCATACTATTTTACAGAGACTTGCTGAAGGATTAAAAGGATTTTCT
 CTTTTGGAAAAGCTTGACTGATTTACACTTATCTATAGTATGCTTTTTTGTGGTGTCTGCTGAATTTAAATAT
 TTATGTGTTTTTCTGTAGGTTGATTTTTTTTGAATCAATATGCAATGTTAAACACTTTTTTAATGTAATCA
 TTTGCATTGGTTAGGAATTCAGAATTCGCCCGGCTCTATTACTGGTCAAGTACATCTTTTCTTTAAATATT
 TAGCCTCCATTATTACAAAAAATTATAAAAAATAAGTTTTTCAAGTCAAGTCAAGGATGACATCACTCCCAATGTTATG
 CAGACATACAGACGGTTGGCATACTGTATAGACTGTATACTCAGTGCAAATATAGCTGCATTTATACCTCAGAG
 GGGCCAAGTGTTAATGCCCATGCCCTCCGTTAAGGGTTGTTGGTTTTACTGGTAGACAGATGTTTTGTGGATTG
 AAAATTATTTTTATGGAATTGCTACAGAGGAGTGCTTTTCTTCTCAATTGTTAGAAGAATTTATGTTAAACTTTA
 AGGTAAGGGTGTA AAAACATTTTTGAGATAAGGTTTTTATTATGTTTATTATTGTTAGAGTGAGTTGCAATGT
 GGAAGAAATGACATTGAAATTCAGTTTTTGAATCCTGTTTCTATTTATAAGTGAAATTTGTGATCTCCTATC
 AACCTTTCATGTTTTACCTGTATAAATGGACATACATGGAACCACTACTGATGAGGGACAGTTGTATGTTTGC
 ATCATATATGCCAGAAAACCTTCCTCTGCTTCCTCTTTTGAAGTTATTTGGTATGTTGTATATATACATAAAA
 TAACTTTTTCAAATATAGTTTAATAACACTTAGAAGTGTTTACTTACCTGGAAAATAATTGCTATGCCGTACATT
 CAGAGTGCCCCCTCCCCTGCAAGGCCTTGCCATGATTAACAAGTAAGTTGTTAGTCTTACAGATAATTATGCA
 TTAACAGTTTAAGATTTAGACCATGGTAATAGTAGTTCTTATTCTCTAAGGTTATATCATATGTAATTTAAAG
 TATTTTTAAGACAAGTTTCCTGTATACCTCTGAAGTTTGTGTTTTGAGTTCATCATGATAGATCTGCTGTTT
 CCTTATAAAAGGCATTTGTTGTGTGAGTTAATGCAAAGTAGCCAAGTCCAGCTATATAGCAGCTTCAGAAACAT
 ACCTGACCAAAAAATTTCCAGTAACCAGGCATGATCAATTTATAGTGGTCGTTTACATCTAATAATTATCAGGA
 CTTTTTTCAGGAGTGGGTTATAAAAACATTCAAGTTGGTCTGACAGTATTTTGTAAAGGATATTTGTTTGTATG
 TTTATTCAGTATACTTACATAAAAATTATTTGCCATCAGCCAAACTCAGTAATCATGACAGCTGTCTGTTGT
 TTTATGAAGTTTATTTCTCAAGAAAATGGGAATAAATTTGGGATTTGTTTCAAGTTTACTAAAGATGCCTAA
 AGCCACAGGTTTTATTGCCTAACTTAAGCCATGACTTTTATAGATATGAGATGACGGGAAGCAGGACGAAATATCG
 GCGTGTGGCTGGAGCCTTCCCCTGGAGGCTGAAAGTGGCTTGTGGTATTATAATGTTTCAAGAGGAA
 GGTGCAGGTACACATGAGTTAGAGAGCTGGTGAGACAGTTGGGAAGTCTTTGTGCTTGTGATCTACTGGACTTT
 TTTTTTGCAGGAAGTGCAATCTCTGGTCCTTCCCTATTTTCTGTTCTGGATGTCAGTGCAGTGCAGTGCCTACTG
 TTTTATCCACTTGGCCACAGACTTTTTCTAACAGCTGCGTATTATTTCTATATACTAATTGCATTGGCAGCATT
 GTGTCTTTGACCTTGTATACTAGCTTGACATAGTGCTGTCTCTGATTTCTAGGCTAGTTACTTGAGATATGAAT
 TTTCCATAGAATATGCACTGATACAACATTACCATTCTTCTATGGAAAGAAAACTTTTGATGATGAAACAATAA
 AGATTTTTAAATATCTATTTTAAAAA

194/330

FIGURE 194

MAGIKALISLSFGGAIGLMFLMLGICALPIYNKYWPLFVLFFYILSPIPYCIARRLVDDTDAM
SNACKELAIFLTGTGIVVSAFGLPIVFARAHLEWGACALVLTGNTVIFATILGFFLVFGSND
DFSWQQW

099997 4904
T06111 22/03/60

[illegible]

FIGURE 196

MDFLLLGLCLYWLLRRPSGVVLCLLGACFQMLPAAPSGCPQLCRCEGRLLYCEALNLTEAPH
NLSGLLGLSLRYNSLSELRAGQFTGLMQLTWLYLDHNNHICSVQGDAFQKLRRVKELTLSSNQ
ITQLPNTTFRPMPNLRSDLSYNKLQALAPDLFHGLRKLTTLHMRANAIQFVPVRIFQDCRS
LKFLDIGYNQLKSLARNSFAGLFKLTELHLEHNDLVKVNFAHFPRLLISLHSLCLRRNKVAIV
VSSLDWVWNLEKMDLSGNEIEYMEPHVFETVPHLQSLQLDSNRLTYIEPRILNSWKSITSIT
LAGNLWDCGRNVCALASWLSNFQGRYDGNLQCASPEYAQGEDVLDVAVYAFHLCEDGAEPTSG
HLLSAVTNRSDLGPPASSATTLADGGEGQHDGTFEPATVALPGGEHAENAVQIHKVVTGTMA
LIFSFLIVVLVLYVSWKCFPASLRQLRQCFVTQRRKQKQKQTMHQMAAMSAQEYYVDYKPNH
IEGALVIINEYGSTCHQQPARECEV

198/330

FIGURE 198

MGVLGRVLLWLQLCALTQAVSKLWVPNTDFDVAANWSQNRTPCAGGAVEFPADKMVSVLVQE
GHAVSDMLLPLDGELVLASGAGFGVSDVGSHLDCGAGEPAVFRDSDRFSWHDPHLWRSGDEA
PGLFFVDAERVPCRHDDVFFPPSASFRVGLGPGASPVRVRSISALGRTFTRDEDLAVFLASR
AGRLRFHGPALSVGPEDCADPSGCVCGNAEAQPWICAALLQP

0989727 44904
"06" "226860"

FIGURE 199

ATCGCATCAATTGGGAGTACCATCTTCCTC**ATG**GGACCAGTGAAACAGCTGAAGCGAATGTT
TGAGCCTACTCGTTTGATTGCAACTATCATGGTGCTGTTGTGTTTTGCACTTACCCTGTGTT
CTGCCTTTTGGTGGCATAACAAGGGACTTGCACTTATCTTCTGCATTTTGCAGTCTTTGGCA
TTGACGTGGTACAGCCTTTCCTTCATAACCATTTGCAAGGGATGCTGTGAAGAAGTGTTCG
CGTGTGTCTTGCA**TAA**ATTCATGGCCAGTTTTATGAAGCTTTGGAAGGCACTATGGACAGAAG
CTGGTGGACAGTTTTGTAACCTATCTTCGAAACCTCTGTCTTACAGACATGTGCCTTTTATCT
TGCAGCAATGTGTTGCTTGCTGATTGCAACATTTGAGGGTTACTTTTGAAGCAACAATACAT
TCTCGAACCTGAATGTCAGTAGCACAGGATGAGAAGTGGGTCTGTATCTTGTGGAGTGGAA
TCTTCCTCATGTACCTGTTTCCTCTCTGGATGTTGTCCCACTGAATCCCATGAATACAAAC
CTATTCAGCAACAGCAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 202

MTSKFILVSFILAALSLSTTFSLQLDQQKVLLVSFDGFRWDYLYKVPTPHFHYIMKYGVHVK
QVTNVFITKTYPNHYTLVTGLFAENHGIVANDMFDPIRNKSFSLDHMNIYDSKFWEEATPIW
ITNQ RAGHTSGAAMWPGTDVKIHKRFPTHYMPYNESVSFEDRVAKIVEWFTSKEPINLGLLY
WEDPDDMGHHLGPDSPLMGPVISDIDKKLGYLIQMLKKAKLWNTLNLIITSDHGMTQCSEER
LIELDQYLDKDHYTLDQSPVAAAILPKEGKFDEVYEALTHAHPNLTVYKKEDVPERWHYKYN
SRIQPIIAVADEGWHILQNKSDDFLLGNHGYDNALADMHPIFLAHGPAFRKNFSKEAMNSTD
LYPLLCHLLNITAMPHNGSFWNVQDLLNSAMPRVVPYTQSTILLPGSVKPAEYDQEGSYPYF
IGVSLGSIIVIVFFVIFIKHLSQIPALQDMHAEIAQPLLQA

Signal Peptide:

amino acids 1-22

Transmembrane Domain:

amino acids 429-452

N-glycosylation sites:

amino acids 101-104, 158-161, 292-295, 329-332, 362-365, 369-372, 382-385, 389-392

Somatomedin B Domain:

amino acids 69-85

Sulfatase protein Region:

amino acids 212-241

FIGURE 203

GGATTTTTGTGATCCGCGATTTCGCTCCACGGGCGGGACCTTTGTAACTGCGGGAGGCCCAG
GACAGGCCCACCCTGCGGGGCGGGAGGCAGCCGGGGTGAGGGAGGTGAAGAAACCAAGACGC
AGAGAGGCCAAGCCCCTTGCCTTGGGTACACAGCCAAAGGAGGCAGAGCCAGAACTCACAA
CCAGATCCAGAGGCAACAGGGAC**ATG**GCCACCTGGGACGAAAAGGCAGTCACCCGCAGGGCC
AAGGTGGCTCCCGCTGAGAGGATGAGCAAGTTCTTAAGGCACTTCACGGTCGTGGGAGACGA
CTACCATGCCTGGAACATCAACTACAAGAAATGGGAGAATGAAGAGGAGGAGGAGGAGGAGG
AGCAGCCACCACCCACACCAGTCTCAGGCGAGGAAGGCAGAGCTGCAGCCCCCTGACGTTGCC
CCTGCCCCCTGGCCCCGCACCCAGGGCCCCCCTTGACTTCAGGGGCATGTTGAGGAACTGTT
CAGCTCCACAGGTTTCAGGTCATCATCATCTGCTTGGTGGTTCTGGATGCCCTCCTGGTGC
TTGCTGAGCTCATCCTGGACCTGAAGATCATCCAGCCCGACAAGAATAACTATGCTGCCATG
GTATTCACCTACATGAGCATCACCATCTTGGTCTTTTTTATGATGGAGATCATCTTTAAATT
ATTTGTCTTCCGCCTGAGTTCTTTCACCACAAGTTTGAGATCCTGGATGCCCCGTCGTGGTGG
TGGTCTCATTCATCCTGGACATTGTCCTCCTGTTCCAGGAGCACCAGTTTGAGGCTCTGGGC
CTGCTGATTCTGCTCCGGCTGTGGCGGGTGGCCCGGATCATCAATGGGATTATCATCTCAGT
TAAGACACGTTCAGAACGGCAACTCTTAAGGTTAAAACAGATGAATGTACAATTGGCCGCCA
AGATTCAACACCTTGAGTTCAGCTGCTCTGAGAAGCCCCCTGGACT**TGA**TGAGTTTGCTGTATC
AACCTGTAAGGAGAAGCTCTCTCCGGATGGCTATGGGAATGAAAGAATCCGACTTCTACTCT
CACACAGCCACCGTGAAAGTCCTGGAGTAAAATGTGCTGTGTACAGAAGAGAGAGAAGGAAG
CAGGCTGGCATGTTCACTGGGCTGGTGTACGACAGAGAACCTGACAGTCACTGGCCAGTTA
TCACTTCAGATTACAAATCACACAGAGCATCTGCCTGTTTTCAATCACAAGAGAACAAAACC
AAAATCTATAAAGATATTCTGAAAATATGACAGAATTTGACAAATAAAAGCATAAACGTGTA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

204/330

FIGURE 204

MATWDEKAVTRRAKVAPAERM SKFLRHFTVVGDDYHAWNIN YKKWENEEEEEEEEQPPPTPV
SGEEGRAAAPDVAPAPGPAPRAPLDFRGMLRKLFSSHRFQV I I ICLVVL DALLVLAELILDL
KIIQPDKN NYAAMVFHYMSITILVFFMMEIIFKLFVFR LSSFTTSLRSWMPVVVVVSFILD I
VLLFQEHQFEALG L L L LRLWRVARIINGI I I SVKTRSERQLLR LKQMN VQLAAKIQHLEFS
CSEKPLD

protein sequence

FIGURE 208

MASLGQILFWSIISIIIIILAGAIALIIGFGISGRHSITVTTVASAGNIGEDGILSCTFEPDI
KLSDIVIQWLKEGVLGLVHEFKEGKDELSEQDEMFRGRTAVFADQVIVGNASLRLKNVQLTD
AGTYKCYIIITSKGKGNANLEYKTGAFSMPEVNVDYNASSETLRCEAPRWFQPTVVASQVD
QGANFSEVSNTSFELNSENVTMKVSVLYNVTINNTYSCMIENDIAKATGDIKVTESEIKRR
SHLQLLSKASLCVSSFFAISWALLPLSPYLMLK

bioRxiv preprint doi: <https://doi.org/10.1101/2020.04.14.240660>; this version posted April 14, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 209

GAATTTGTAGAAGACAGCGGCGTTGCC**ATG**GCGGCGTCTCTGGGGCAGGTGTTGGCTCTGGT
 GCTGGTGGCCGCTCTGTGGGGTGGCACGCAGCCGCTGCTGAAGCGGGCCTCCGCCGGCCTGC
 AGCGGGTTTCATGAGCCGACCTGGGCCCAGCAGTTGCTACAGGAGATGAAGACCCTCTTCTTG
 AATACTGAGTACCTGATGCCCTTTCTCCTCAACCAGTGTGGATCCCTTCTCTATTACCTCAC
 CTTGGCATCGACAGATCTGACCCTGGCTGTGCCCATCTGTAACCTCTCTGGCTATCATCTTCA
 CACTGATTGTTGGGAAGGCCCTTGGAGAAGATATTGGTGGAAAACGTAAGTTAGACTACTGC
 GAGTGCGGGACGCAGCTCTGTGGATCTCGACATACCTGTGTTAGTTCCCTTCCCAGAACCCAT
 CTCCCCAGAGTGGGTGAGGACACGGCCTTTTCCCATCCTGCCCTTTCTCTGCAGCTGTTTT
 GCTTCCCTGTGGCCATCAGAGTTCCCTTCCCCTGGACAGTCTGGAGAAAGACAGAGGCTGGG
 GTTTGGGAT**TGA**AAGACCAGACCCCATCTGAGCCCTTCCCTCCAGCCCTGTACCAGCTCCTACT
 GGCATGGCTGAGCTCAGACCCTCCTGATTTCTGCCTATTATCCCAGGAGCAGTTGCTGGCAT
 GGTGCTCACCGTGATAGGAATTTCACTCTGCATCACAAGCTCAGTGAGTAAGACCCAGGGGC
 AACAGTCTACCCTTTGAGTGGGCCGAACCCACTTCCAGCTCTGCTGCCTCCAGGAAGCCCCT
 GGGCCATGAAGTGCTGGCAGTGAGCGGATGGACCTAGCACTTCCCCTCTCTGGCCTTAGCTT
 CCTCCTCTCTTATGGGGATAACAGCTACCTCATGGATCACAATAAGAGAACAAGAGTGAAAG
 AGTTTTGTAACTTCAAGTGCTGTTCACTGCTGCGGGGATTTAGCACAGGAGACTCTACGCTCA
 CCCTCAGCAACCTTTCTGCCCCAGCAGCTCTCTTCCCTGCTAACATCTCAGGCTCCCAGCCCA
 GCCACCATTACTGTGGCCTGATCTGGACTATCATGGTGGCAGGTTCCATGGACTGCAGAACT
 CCAGCTGCATGGAAAGGGCCAGCTGCAGACTTTGAGCCAGAAATGCAAACGGGAGGCCTCTG
 GGACTCAGTCAGAGCGCTTTGGCTGAATGAGGGGTGGAACCGAGGGAAGAAGGTGCGTCGGA
 GTGGCAGATGCAGGAAATGAGCTGTCTATTAGCCTTGCCCTGCCCCACCCATGAGGTAGGCAG
 AAATCCTCACTGCCAGCCCCTCTTAAACAGGTAGAGAGCTGTGAGCCCCAGCCCCACCTGAC
 TCCAGCACACCTGGCGAGTAGTAGCTGTCAATAAATCTATGTAAACAGACAAAAAAAAAAAAA
 AA

212/330

FIGURE 212

MKGILVAGITAVLVAAVESLSCVQCNSWEKSCVNSIASECPSHANTSCISSSASSSLETPVR
LYQNMFCSAENCSEETHITAFTVHVSAEEHFHFVSQCCQGKECSNTSDALDPPLKNVSSNAE
CPACYESNGTSCRGKPWKCYEEEQCVFLVAELKNDIESKSLVLKGCSNVSNATCQFLSGENK
TLGGVIFRKFECANVNSLTPTSAPTTSNHNVGSKASLYLLALASLLLRGLLP

bioRxiv preprint doi: <https://doi.org/10.1101/2021.04.28.440000>; this version posted May 3, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 213

GGCCTCGGTTCAAACGACCCGGTGGGTCTACAGCGGAAGGGAGGGAGCGAAGGTAGGAGGCA
GGGCTTGCCCTCACTGGCCACCCTCCCAACCCCAAGAGCCCAGCCCC**ATG**GTCCCCGCCGCCG
GCGCGCTGCTGTGGGTCTGCTGCTGAATCTGGGTCCCCGGGCGGCGGGGGCCCAAGGCCTG
ACCCAGACTCCGACCGAAATGCAGCGGGTCAGTTTACGCTTTGGGGGCCCCATGACCCGCAG
CTACCGGAGCACCGCCCGGACTGGTCTTCCCCGGAAGACAAGGATAATCCTAGAGGACGAGA
ATGATGCCATGGCCGACGCCGACCGCCTGGCTGGACCAGCGGCTGCCGAGCTCTTGGCCGCC
ACGGTGTCCACCGGCTTTAGCCGGTCGTCCGCCATTAACGAGGAGGATGGGTCTTCAGAAGA
GGGGGTGTGATTAATGCCGGAAGGATAGCACCAGCAGAGAGCTTCCCAGTGCGACTCCCA
ATACAGCGGGGAGTTCCAGCACGAGGTTTATAGCCAATAGTCAGGAGCCTGAAATCAGGCTG
ACTTCAAGCCTGCCGCGCTCCCCGGGAGGTCTACTGAGGACCTGCCAGGCTCGCAGGCCAC
CCTGAGCCAGTGGTCCACACCTGGGTCTACCCCGAGCCGGTGGCCGTACCCTCACCCACAG
CCATGCCATCTCCTGAGGATCTGCGGCTGGTGCTGATGCCCTGGGGCCCGTGGCACTGCCAC
TGCAAGTCGGGCACCATGAGCCGGAGCCGGTCTGGGAAGCTGCACGGCCTTTCGGGGCGCCT
TCGAGTTGGGGCGCTGAGCCAGCTCCGCACGGAGCACAAGCCTTGACCTATCAACAATGTC
CCTGCAACCGACTTCGGGAAGAGTGCCCCCTGGACACAAGTCTCTGTACTGACACCAACTGT
GCCTCTCAGAGCACCACCAGTACCAGGACCACCACTACCCCTTCCCCACCATCCACCTCAG
AAGCAGTCCCAGCCTGCCACCCGCCAGCCCTGCCAGCCCTGGCTTTTTGGAAACGGGTCA
GGATTGGCCTGGAGGATATTTGGAATAGCCTCTCTTCAGTGTTACAGAGATGCAACCAATA
GACAGAAACCAGAGG**TAA**TGGCCACTTCATCCACATGAGGAGATGTCAGTATCTCAACCTCT
CTTGCCCTTTCAATCCTAGCACCCACTAGATATTTTTAGTACAGAAAAACAAAACCTGGAAAA
CACAA

FIGURE 214

MVPAAGALLWVLLLNLGPRAGAQLTQTPTEMQRVSLRFGGPMTRSYRSTARTGLPRKTRI
ILEDENDAMADADRLAGPAAELLAATVSTGFSRSSAINEEDGSSEEGVVINAGKDSTSREL
PSATPNTAGSSSTRFIANSQEPEIRLTSSLPRSPGRSTEDLPGSQATLSQWSTPGSTPSRWP
SPSPTAMPSPEDLRLVLMWPWGPWHCHCKSGTMSRSRSGKLHGLSGRLRVGALSQLRTEHKPC
TYQQCPCNRLREECPLDTSLCDTNCASQSTTSTRTTTTPFPTIHLRSSPSLPPASPCPALA
FWKRVIRIGLEDIWNSLSSVFTEMQPIDRNQR

214/330

FIGURE 216

MAGLAARLVLLAGAAALASGSQGDREPVYRDCVLQCEEQNCSGGALNHFRSRQPIYMSLAGW
TCRDDCKYECMWVTVGLYLQEGHKVPQFHGKWPFSSRFLFFQEPASAVASFLNGLASLVMLCR
YRTFVPASSPMYHTCVAFAWVSLNAFWSTVFHTRDTDLTEKMDYFCASTVILHSIYLCCVR
TVGLQHPAVVSAFRALLLLMLTVHVSYSLSLIRFDYGYNLVANVAIGLVNVVWWLAWCLWNQR
RLPHVRKCVVVVLLQGLSLELLDFPPLFWVLDAAHAIWHISTIPVHVLFFSFLEDDSLYLL
KESEDKFKLD

Important features:**Signal peptide:**

amino acids 1-20

Transmembrane domains:

amino acids 105-123, 138-156, 169-185, 193-209, 221-240, 256-272

N-glycosylation site.

amino acids 40-44

N-myristoylation site.

amino acids 43-49

CUB domain proteins profile.

amino acids 285-302

Amiloride-sensitive sodium channels proteins.

amino acids 162-186

FIGURE 217

GGCCGCCTGGAATTGTGGGAGTTGTGTCTGCCACTCGGCTGCCGGAGGCCGAAGGTCCGTGA
CT**ATG**GCTCCCCAGAGCCTGCCTTCATCTAGGATGGCTCCTCTGGGCATGCTGCTTGGGCTG
CTGATGGCCGCCTGCTTCACCTTCTGCCTCAGTCATCAGAACCTGAAGGAGTTTGCCCTGAC
CAACCCAGAGAAGAGCAGCACCAAAGAAACGGAGAGAAAAGAAACCAAAGCCGAGGAGGAGC
TGGATGCCGAAGTCCTGGAGGTGTTCCACCCGACGCATGAGTGGCAGGCCCTTCAGCCAGGG
CAGGCTGTCCCTGCAGGATCCCACGTACGGCTGAATCTTCAGACTGGGGAAAGAGAGGCAAA
ACTCCAATATGAGGACAAGTTCCGAAATAATTTGAAAGGCAAAAGGCTGGATATCAACACCA
ACACCTACACATCTCAGGATCTCAAGAGTGCCTGGCAAAATTCAAGGAGGGGGCAGAGATG
GAGAGTTCAAAGGAAGACAAGGCAAGGCAGGCTGAGGTAAAGCGGCTCTTCCGCCCCATTGA
GGAAGTGAAGAAAGACTTTTGATGAGCTGAATGTTGTCATTGAGACTGACATGCAGATCATGG
TACGGCTGATCAACAAGTTCAATAGTTCCAGCTCCAGTTTGGAAGAGAAGATTGCTGCGCTC
TTTGATCTTGAATATTATGTCCATCAGATGGACAATGCGCAGGACCTGCTTTCCTTTGGTGG
TCTTCAAGTGGTGATCAATGGGCTGAACAGCACAGAGCCCCCTCGTGAAGGAGTATGCTGCGT
TTGTGCTGGGCGCTGCCTTTTCCAGCAACCCCAAGGTCCAGGTGGAGGCCATCGAAGGGGGA
GCCCTGCAGAAGCTGCTGGTCATCCTGGCCACGGAGCAGCCGCTCACTGCAAAGAAGAAGGT
CCTGTTTGCCTGCTGCTCCCTGCTGCGCCACTTCCCCTATGCCAGCGGCAGTTCCTGAAGC
TCGGGGGGCTGCAGGTCCTGAGGACCCTGGTGCAGGAGAAGGGCACGGAGGTGCTCGCCGTG
CGCGTGGTCACACTGCTCTACGACCTGGTCACGGAGAAGATGTTCCGCCGAGGAGGAGGCTGA
GCTGACCCAGGAGATGTCCCCAGAGAAGCTGCAGCAGTATCGCCAGGTACACCTCCTGCCAG
GCCTGTGGGAACAGGGCTGGTGCAGATCACGGCCCACCTCCTGGCGCTGCCCCAGCATGAT
GCCCCGTGAGAAGGTGCTGCAGACACTGGGCGTCCTCCTGACCACCTGCCGGGACCGCTACCG
TCAGGACCCCCAGCTCGGCAGGACACTGGCCAGCCTGCAGGCTGAGTACCAGGTGCTGGCCA
GCCTGGAGCTGCAGGATGGTGAGGACGAGGGCTACTTCCAGGAGCTGCTGGGCTCTGTCAAC
AGCTTGCTGAAGGAGCTGAGAT**TGA**AGGCCCCACACCAGGACTGGACTGGGATGCCGCTAGTGA
GGCTGAGGGGTGCCAGCGTGGGTGGGCTTCTCAGGCAGGAGGACATCTTGGCAGTGCTGGCT
TGGCCATTAAATGGAACCTGAAGGCCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AA

FIGURE 218

MAPQSLPSSRMAPLGMLLGLLMAACFTFCLSHQNLKEFALTNPEKSSTKETERKETKAEHEEL
 DAEVLEVFHPTHEWQALQPGQAVPAGSHVRLNLQTGEREAKLQYEDKFRNNLKGKRLDINTN
 TYTSQDLKSALAKFKEGAEMESSKEDKARQAEVKRLFRPIEELKKDFDELNVVIETDMQIMV
 RLINKFNSSSSSLEEKIAALFDLEYVYVHQMDNAQDLLSFQGLQVVINGLNSTEPLVKEYAAF
 VLGAAFSSNPVKVQVEAIEGGALQKLLVILATEQPLTAKKKVLFALCSLLRHFPYAQRQFLKL
 GGLQVLRRTLTVQEKGTENVLAVRVVTLTYDLVTEKMFEEEEAELETQEMSPEKLQQYRQVHLLPG
 LWEQGWCEITAHLLALPEHDAREKVLQTLGVLLTTCRDYRQDPQLGRTLASLQAEYQVLAS
 LELQDGEDEGYFQELLGSVNSLLKELR

Important features:**Signal peptide:**

amino acids 1-29

Hypothetical YJL126w/YLR351c/yhcX family protein.

amino acids 364-373

N-glycosylation site.

amino acids 193-197, 236-240

N-myristoylation site.

amino acids 15-21, 19-25, 234-240, 251-257, 402-408, 451-457

Homologous region SLS1 protein.

amino acids 68-340

FIGURE 223

NGTTGGAGAAGTGGCGCGGACNTTCATTTGGGGTTTCGGTTTCCCCCCTTTCCCTTTCCCCG
GGGTCTGGGGTGACATTGCACGGGCCCCTCGTGGGGTCGCGTTGCCACCCCACGCGGACTCC
CCAGNTGGNGCGCCCTTCCCATTTGCCTGTCCTGGTCAGGCCCCCACCCTTCCCACNTG
ACCAGCCATGGGGGCTGCGGTGTTTTTCGGCTGCACTTTCGTCGCGTTCGGCCCGGCCTTCG
CGCTTTTCTTGATCACTGTGGCTGGGGACCCGCTTCGCGTTATCATCCTGGTCGCAGGGGCA
TTTTTCTGGCTGGTCTCCCTGCTCCTGGCCTCTGTGGTCTGGTTCATCTTGGTCCATGTGAC
CGACCGGTCAGATGCCCCGGCTCCAGTACGGCCTCCTGATTTTTTGGTGCTGCTGTCTCTGTCC
TTCTACAGGAGGTGTTCCGCTTTGCCTACTACAAGCTGCTTAAGAAGGCAGATGAGGGGTTA
GCATCGCTGAGTGAGGACGGAAGATCACCCATCTCCATCCGCCAGATGGCCTATGTTTCTGG
TCTCTCCTTCGGTATCATCAGTGGTGTCTTCTCTGTTATCAATATTTTGGCTGATGCACTTG
GGCCAGGTGTGGTTGGGATCCATGGAGACTCACCC

FIGURE 224

GTAAAAGAAAGTGGCCGGACCTTCATTGGGGTTTCGGTTCCCCCTTTCCCNTTCCCCGGGG
TCTGGGGGTGACATTGCACCGCGCCCNTCGTGGGGTCGCGTTGCCACCCACGCGGACTCCC
CAGNTGGCGCGCCCCCTCCCATTTCCTGTCCTGGTCAGGCCCCACCCCCCTTCCCACCTGA
CCAGCCATGGGGGCTGCGGTGTTTTTCGGGCTGCACTTTCGTCGCGTTCGGGCCCCGGCCTTC
GCGCTTTTCTTGATCACTGTGGCTGGGGACCGCTTCGCGTTATCATCCTGGTCGCAGGGGC
ATTTTTCTGGCTGGTCTCCCTGCTCCTGGCCTCTGTGGTCTGGTTCATCTTGGTCCATGTGA
CCGACCGGTGAGATGCCCGGCTCCAGTACGGCCTCCTGATTTTTGGTGCTGCTGTCTCTGTC
CTTCTACAGGAGGTGTTCCGCTTTCCTACTACAAGCTGCTTAAGAAGGCAGATGAGGGGTT
AGCATCGCTGAGTGAGGACGGAAGATCACCCATCTCCATCCGCCAGATGGCCTATGTTTCTG
GTCTCTCCTTCGGTATCATCAGTGGTGTCTTCTCTGTTATCAATATTTTGGCTGATGCACTT
GGGCCAGGTGTGGTTGGGATCCATGGAGAC

FIGURE 226

MATARPPWMWVLCALITALLLGVT EHVLANNDVSCDHPSNTVPSGSNQDLGAGAGEDARSDD
SSSRIINGSDCDMHTQPWQAALLLRPNQLYCGAVLVHPQWLLTAAHCRKKVFRVRLGHYSLS
PVYESGQQMFQGVKSI PHPGYSHPGHSNDLMLIKLNRRIRPTKDVRPINVSSHCP SAGTKCL
VSGWGTTKSPQVHF PKVLQCLNISVLSQKRCE DAYPRQIDDTMFCAGDKAGR DSCQGD SGGP
VVCNGSLQGLVSWGDYPCARPNRPGVYTNLCKFTKWIQETIQANS

FIGURE 229

GCAGTCAGAGACTTCCCCTGCCCCCTCGCTGGGAAAGAACATTAGGAATGCCTTTTAGTGCCT
 TGCTTCCTGAACTAGCTCACAGTAGCCCGGCGGCCAGGGCAATCCGACCACATTTCACTCT
 CACCGCTGTAGGAATCCAGATGCAGGCCAAGTACAGCAGCACGAGGGACATGCTGGATGATG
 ATGGGGACACCACCATGAGCCTGCATTCTCAAGCCTCTGCCACAACTCGGCATCCAGAGCCC
 CGGCGCACAGAGCACAGGGCTCCCTCTTCAACGTGGCGACCAGTGGCCCTGACCCTGCTGAC
 TTTGTGCTTGGTGCTGCTGATAGGGCTGGCAGCCCTGGGGCTTTTGTTTTTTTCAGTACTACC
 AGCTCTCCAATACTGGTCAAGACACCATTTCTCAAATGGAAGAAAGATTAGGAAATACGTCC
 CAAGAGTTGCAATCTCTTCAAGTCCAGAATATAAAGCTTGCAGGAAGTCTGCAGCATGTGGC
 TGAAAACTCTGTCGTGAGCTGTATAACAAAGCTGGAGCACACAGGTGCAGCCCTTGTACAG
 AACAATGGAATGGCATGGAGACAATTGCTACCAGTTCTATAAAGACAGCAAAAGTTGGGAG
 GACTGTAAATATTTCTGCCTTAGTGAAACTCTACCATGCTGAAGATAAAACAAACAAGAAGA
 CCTGGAATTTGCCGCGTCTCAGAGCTACTCTGAGTTTTTCTACTCTTATTGGACAGGGCTTT
 TGCGCCCTGACAGTGGCAAGGCCTGGCTGTGGATGGATGGAACCCCTTTCACTTCTGAACTG
 TTCCATATTATAATAGATGTCACCAGCCCAAGAAGCAGAGACTGTGTGGCCATCCTCAATGG
 GATGATCTTCTCAAAGGACTGCAAAGAATTGAAGCGTTGTGTCTGTGAGAGAAGGGCAGGAA
 TGGTGAAGCCAGAGAGCCTCCATGTCCCCCTGAAACATTAGGCGAAGGTGACTGATTTCGCC
 CTCTGCAACTACAAATAGCAGAGTGAGCCAGGCGGTGCCAAAGCAAGGGCTAGTTGAGACAT
 TGGGAAATGGAACATAATCAGGAAAGACTATCTCTCTGACTAGTACAAAATGGGTTCTCGTG
 TTTCTGTTCAGGATCACCAGCATTTCTGAGCTTGGGTTTATGCACGTATTTAACAGTCACA
 AGAAGTCTTATTTACATGCCACCAACCAACCTCAGAAACCCATAATGTCATCTGCCTTCTTG
 GCTTAGAGATAACTTTTAGCTCTCTTTCTTCTCAATGTCTAATATCACCTCCCTGTTTTTCAT
 GTCTTCCTTACACTTGGTGGAATAAGAACTTTTTGAAGTAGAGGAAATACATTGAGGTAAC
 ATCCTTTTCTCTGACAGTCAAGTAGTCCATCAGAAATTGGCAGTCACTTCCCAGATTGTACC
 AGCAAATACACAAGGAATTCTTTTTGTTTGTTCAGTTTCATACTAGTCCCTTCCCAATCCAT
 CAGTAAAGACCCCATCTGCCTTGTCCATGCCGTTTCCCAACAGGGATGTCACTTGATATGAG
 AATCTCAAATCTCAATGCCTTATAAGCATTCCTTCCTGTGTCCATTAAGACTCTGATAATTG
 TCTCCCCTCCATAGGAATTTCTCCCAGGAAAGAAATATATCCCCATCTCCGTTTCATATCAG
 AACTACCGTCCCCGATATTCCCTTCAGAGAGATTAAAGACCAGAAAAAAGTGAGCCTCTTCA
 TCTGCACCTGTAATAGTTTCAGTTCCTATTTTCTTCCATTGACCCATATTTATACCTTTTCAG
 GTACTGAAGATTTAATAATAATAAATGTAAATACTGTGAAAAA

FIGURE 230

MQAKYSSTRDMLDDDGDTTMSLHSQASATTRHPEPRRTEHRAPSSTWRPVALTLLTLCLVLL
IGLAALGLLFFQYYQLSNTGQDTISQMEERLGNTSQELQSLQVQNIKLAGSLQHVAEKLCRE
LYNKAGAHRCSPCTEQWKWHGDNCYQFYKDSKSWEDCKYFCLSENSTMLKINKQEDLEFAAS
QSYSEFFYSYWTGLLRPDSGKAWLWMDGTPFTSELFHIIIDVTSPRSRDCVAILNGMIFSKD
CKELKRCVCERRAGMVKPESLHVPPETLGED

FIGURE 233

MRASLLLSVLRPAGPVAVGISLGFTLSLLSVTWVEEPCGPGPPQPGDSELP PRGNTNAARRP
NSVQPGAEREKPGAGEGAGENWEPRVLPYHPAQPGQAACKAVRTRYISTELGIRQLLVAVL
TSQTTLPTLGVAVNRTLGHRLERVVFLTGARGRRAPPGMAVVTLGEERPIGHLHLALRHLL
QHGDDEFDWFFLVPDTTYTEAHGLARLTGHLSLASAAHLYLGRPQDFIGGEPTPGRYCHGGFG
VLLSRMLLQQLRPHLEGCRNDIVSARPDEWLGRCIL DATGVGCTGDHEGVHYSHLELSPGEP
VQEGDPHFRSALTAHPVRDPVHMYQLHKAFARAELE RTYQEIQELQWEIQNTSHLAVDGDRA
AAWPVGIPAPSRPASRFEVLRWDYFTEQHAFSCADGSPRCPLRGADRADVADVLGTALEELN
RRYHPALRLQKQQLVNGYRRFDPARGMEYTLDLQLEALTPQGRRPLTRRVQLLRPLSRVEI
LPVPYVTEASRLTVLLPLAAAERDLAPGFLEAFATAALEPGDAAAALTLLLLYEPRQAQRVA
HADVFAPVKAHVAELERRFPGARVPWLSVQTAAPSPLRLMDLLSKKHPLDTLFLLAGPDTVL
TPDFLNRCRMHAISGWQAFFPMHFQAFHPGVAPPQGP GPPELGRDTGRFDRQAASEACFYNS
DYVAARGRLAAASEQEEELLES LDVYELFLHFSSLHVLRAVEPALLQRYRAQTCSARLSEDL
YHRCLQSVLEGLGSRTQLAMLLFEQE QGNST

235/330

FIGURE 235

MAAGLFGLSARLLAAAATRGLPAARVRWESSFSRTVVAPSAVAGKRPPEPTTPWQEDPEPE
DENLYEKNPDSHGYDKDPVLDVWNMRLVFFFGVSIILVLGSTFVAYLPDYRMKEWSRREAER
LVKYREANGLPIMESNCFDPSKIQLPEDE

bioRxiv preprint doi: <https://doi.org/10.1101/2020.03.10.332660>; this version posted March 10, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 238

MPLALLVLLLLGPGGWCLAEPPrDSLREELVITPLPSGDVAATFQFRTRWDSELQREGVSHY
RLFPKALGQLISKYSLRELHLSFTQGFWRTYWGPPFLQAPSGAELWVWFQDTVTDVDSWK
ELSNVLSGIFCASLNFIDSTNTVTPTASFkPLGLANDTDHYFLRYAVLPREVVCTENLTPWK
KLLPCSSKAGLSVLLKADRLFHTSYHSQAVHIRPVCRNARCTSSISWELRQTLsvVFDafITG
QGKDWslFRMFsRTLTEPCPLASESRVYVDITTYNQDNETLEVHPPPTTTYQDVILGTRKT
YAIYDLLDTAMINNSRNLNIQLKWKRPPENEAPPVPFLHAQRYVSGYGLQKGELSTLLYNTH
PYRAFPVLLLDTPWYLRlyVHTLTITSKGKENKPSYIHYQPAQDRLQPHLLEMLIQLPANS
VTKVSIQFERALLKWTEYTPDPNHGFYVSPSVLSALVPSMVAAKpVDweESPLFNSLFPVSD
GSNYFVRLYTEPLLvNLPTPDFSMPYNVICLTCTVVAVCYGSFYNLLTRTFHIEEPRTGGLA
KRLANLIRRARGVPPL

FIGURE 239

CAACATGGGGTCCAGCAGCTTCTTGGTCCTCATGGTGTCTCTCGTTCTTGTGACCCTGGTGG
CTGTGGAAGGAGTTAAAGAGGGTATAGAGAAAGCAGGGGTTTGCCCAGCTGACAACGTACGC
TGCTTCAAGTCCGATCCTCCCCAGTGTCACACAGACCAGGACTGTCTGGGGGAAAGGAAGTG
TTGTTACCTGCACTGTGGCTTCAAGTGTGTGATTCTGTGAAGGAACTGGAAGAAGGAGGAA
ACAAGGATGAAGATGTGTCAAGGCCATACCCTGAGCCAGGATGGGAGGCCAAGTGTCCAGGC
TCCTCCTCTACCAGGTGTCCTCAGAATGATGCTGGGTCCTTTCTACCTCTGGGGGTCACTC
TCACTTGGCACCTGCCCCTGAGGGTCCTGAGACTTGGAATATGGAAGAAGCAATACCCAACC
CCACCAAAGAAAACCTGAGCTTGAAGTCCTTTTCCCCAAAAGAGGGAAGAGTCACAAAAG
TCCAGACCCCAGGGACGGTACTTTCCCTCTCTACCTGGTGCTCCTCCCTAATGCTCATGAAT
GGACCCCTCATGAATGAAACCAGTGCCCTTATAAGAGACCCCAAAGAGCTGCCTTGCCCTTC
TGCAATGTGTGATCACAGCTAGAAGGCACTGTCAGAGAAGAGAACTGGTCCTCACCAGATG
CTGAATCTGCTGGTGCCTTGATCTTGGACTTCCCAGCCTCTAGAACTGTAAGAAATAAATAT
TTGCTGTTTATAATCCAA

240/330

FIGURE 240

MGSSSFLVLMVSLVLVTLVAVEGVKEGIEKAGVCPADNVRCFKSDPPQCHTDQDCLGERKCC
YLHCGFKCVIPVKELEEGGNKDEDVSRPYPEPGWEAKCPGSSSTRCPQK

Signal sequence:

amino acids 1-19

N-myristoylation sites:

amino acids 23-29, 27-33, 32-38, 102-108

WAP-type 'four-disulfide core' domain signature:

amino acids 49-63

bioRxiv preprint doi: <https://doi.org/10.1101/2021.04.14.438888>; this version posted April 14, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

FIGURE 242

MRSCLWRCRHLSQGVQWSLLLVFVFFLFALPSFIKEPQTKPSRHQRTENIKERSLQSLAKP
 KSQAPTRARRTTIYAEPA PENNALNTQTQPKAHTTGDRGKEANQAPPEEQDKVPHTAQRAAW
 KSPEKEKTMVNTLS PRGQDAGMASGRTEAQSWKSQDTKTTQGNNGQTRKLTASRTVSEKHQG
 KAATTAKTLIPKSQHRMLAPTGA VSTRTRQKGVTTAVIPPKEKKPQATPPPAPFQSPTTQRN
 QRLKAANFKSEPRWDFEEKYSFEIGGLQTTCPDSVKIKASKSLWLQKLFLPNLTLFLDSRHF
 NQSEWDRLEHFAPPPFGFMELNYSLVQKVVTFRFPVPQQQLLLASLPAGSLRCITCAVVGNGG
 ILNNSHMGQEIDSHDYVFR LSGALIKGYEQDVGTRTSFYGF TAFSLTQSL LILGNRGFKNVP
 LGKDVRYLHFLEGRDYE WLEALLMNQTVMSKNLFWFRHRPQEAFREALHMDRYLLLHPDFL
 RYMKNRFLRSKTL DGAHWRIYRPTTGALLLLTALQLCDQVSAYGFITEGHERFSDHYYDTSW
 KRLIFYINHDFKLEREVWKRLHDEGIIRLYQRPGPGTAKAKN

Cytoplasmic Domain:

amino acids 1-10

Type II Transmembrane Domain:

amino acids 11-35

Lumenal catalytic Domain:

amino acids 36-600

Ribonucleotide Reductase small subunit Signature:

amino acids 481-496

N-glycosylation Sites:

amino acids 300-303, 311-314, 331-334, 375-378, 460-463

244/330

FIGURE 244

MRGPGHPLLLGLLLVLGPSPEQRVEIVPRDLRMKDKFLKHLTGPLYFSPKCSKHFRLYHNT
RDCTIPAYYKRCARLLTRLAVSPVCMEDK

244/330

FIGURE 245

GGGCTGGGCCCCGCCGCAGCTCCAGCTGGCCGGCTTGGTCCTGCGGTCCCTTCTCTGGGAGG
CCCGACCCCGGCCGCGCCCAGCCCCACCATGCCACCCGCGGGGCTCCGCCGGGCCGCGCCG
CTCACCGCAATCGCTCTGTTGGTGCTGGGGGCTCCCCTGGTGCTGGCCGGCGAGGACTGCCT
GTGGTACCTGGACCGGAATGGCTCCTGGCATCCGGGGTTTAACTGCGAGTTCTTCACCTTCT
GCTGCGGGACCTGCTACCATCGGTACTGCTGCAGGGACCTGACCTTGCTTATCACCGAGAGG
CAGCAGAAGCACTGCCTGGCCTTCAGCCCCAAGACCATAGCAGGCATCGCCTCAGCTGTGAT
CCTCTTTGTTGCTGTGGTTGCCACCACCATCTGCTGCTTCCTCTGTTTCTGTTGCTACCTGT
ACCGCCGGCGCCAGCAGCTCCAGAGCCCATTGAAGGCCAGGAGATTCCAATGACAGGCATC
CCAGTGCAGCCAGTATACCCATACCCCCAGGACCCCCAAAGCTGGCCCTGCACCCCCACAGCC
TGGCTTCATGTACCCACCTAGTGGTCCTGCTCCCCAATATCCACTCTACCCAGCTGGGCCCC
CAGTCTACAACCCTGCAGCTCCTCCTCCCTATATGCCACCACAGCCCTCTTACCCGGGAGCC
TGAGGAACCAGCCATGTCTCTGCTGCCCCCTTCAGTGATGCCAACCTTGGGAGATGCCCTCAT
CCTGTACCTGCATCTGGTCCTGGGGGTGGCAGGAGTCCTCCAGCCACCAGGCCCCAGACCAA
GCCAAGCCCCTGGGCCCTACTGGGGACAGAGCCCCAGGGAAGTGGAACAGGAGCTGAACTAGA
ACTATGAGGGGTTGGGGGGAGGGCTTGGAATTATGGGCTATTTTTTACTGGGGGCAAGGGAGG
GAGATGACAGCCTGGGTCACAGTGCCTGTTTTCAAATAGTCCCTCTGCTCCCAAGATCCCAG
CCAGGAAGGCTGGGGCCCTACTGTTTGTCCCCTCTGGGCTGGGGTGGGGGGAGGGAGGAGGT
TCCGTCAGCAGCTGGCAGTAGCCCTCCTCTCTGGCTGCCCCACTGGCCACATCTCTGGCCTG
CTAGATTAAAGCTGTAAAGACAAAA

FIGURE 246

MPPAGLRRAAPLTAIALLLVLGAPLVLAGEDCLWYLDNRNGSWHPGFNCEFFTFC CGTCYHRYC
CRDLTLLITERQQKHCLAFSPKTIAGIASAVILFVAVVATTICCF LCSCCYLYRRRQQQLQSP
FEGQEIPMTGIPVQPVPYPYPQDPKAGPAPPQPGFMYPPSGPAPQYPLYPAGPPVYNPAAPP
YMPPQPSYPGA

Transmembrane Domains:

amino acids 10-28, 85-110

N-glycosylation Site:

amino acids 38-41

N-myristoylation Sites:

amino acids 5-10, 88-93

FIGURE 248

MVFLPLKWSLATMSFLLSSLLALLTVSTPSWCQSTEASPKRSDGTPFPWNKIRLPEYVIPVH
 YDLLIHANLTTLTFWGTTKVEITASQPTSTIILHSHHLQISRATLRKGAGERLSEEPLQVLE
 HPPQEQIALLAPEPLLVLGPYTVVIHYAGNLSETFHGFYKSTYRTKEGELRILASTQFEPTA
 ARMAFPCFDEPAFKASF SIKIRREPRHLAISNMPLVKSVTVAEGLIEDHFDVTVKMSTYLVA
 FIISDFESVSKITKSGVKVSVYAVPDKINQADYALDAAVTLLLEFYEDYFSIPYPLPKQDLAA
 IPDFQSGAMENWGLTTYRESALLFDAEKSSASSKLGITVTVAHELHQWFGNLVTMEWWNDL
 WLNEGFAKFMEFVS SVTHPELVGDYFFGKCFDAMEVDALNSSHPVSTPVENPAQIREMFD
 DVSYDKGACILNMLREYLSADAFKSGIVQYLQKHSYKNTKNEDLWDSMASICPTDGVKGMDG
 FCSRSQHSSSSSHWHQEGVDVKTMMNTWTLQRGFPLITITVRGRNVHMKQEHYMKGSDGAPD
 TGYLWHVPLTFITSKSNMVHRFLKTKTDVLILPEEVEWIKFNVGMNGYYIVHYEDDGWDSL
 TGLLKGTHTAVSSNDRASLINNAFQLV SIGKLSIEKALDLSLYLKHETEIMPVFQGLNELIP
 MYKLMEKRD MNEVETQFKAF LIRLLRD LIDKQTWTDEGSVSEQMLRSELLLLACVHNYQPCV
 QRAEGYFRKWKESNGNLSLPVDVTLAVFAVGAQSTEGWDFLYSKYQFSLSSTEKSQIEFALC
 RTQNKEKLQWLLDESFKGDKIKTQEFPPQILT LIGRNPVGYPLAWQFLRKNWNKLVQKFELGS
 SSIAHVMVGTTNQFSTRTRLEEVKGFFSSLKENGSQLRCVQQTIETIEENIGWMDKNFDKIR
 VWLQSEKLERM

Signal peptide:

amino acids 1-34

N-glycosylation sites:

amino acids 70-74, 154-158, 414-418, 760-764, 901-905

Neutral zinc metallopeptidases, zinc-binding region signature:

amino acids 350-360

FIGURE 249

CAGCCACAGACGGGTCATGAGCGCGGTATTACTGCTGGCCCTCCTGGGGTTCATCCTCCCAC
TGCCAGGAGTGCAGGCGCTGCTCTGCCAGTTTGGGACAGTTCAGCATGTGTGGAAGGTGTCC
GACCTACCCCGGCAATGGACCCCTAAGAACACCAGCTGCGACAGCGGCTTGGGGTGCCAGGA
CACGTTGATGCTCATTGAGAGCGGACCCCAAGTGAGCCTGGTGCTCTCCAAGGGCTGCACGG
AGGCCAAGGACCAGGAGCCCCGCGTCACTGAGCACCGGATGGGCCCCGGCCTCTCCCTGATC
TCCTACACCTTCGTGTGCCGCCAGGAGGACTTCTGCAACAACCTCGTTAACTCCCTCCCGCT
TTGGGCCCCACAGCCCCCAGCAGACCCAGGATCCTTGAGGTGCCAGTCTGCTTGTCTATGG
AAGGCTGTCTGGAGGGGACAACAGAAGAGATCTGCCCCAAGGGGACCACACACTGTTATGAT
GGCCTCCTCAGGCTCAGGGGAGGAGGCATCTTCTCCAATCTGAGAGTCCAGGGATGCATGCC
CCAGCCAGGTTGCAACCTGCTCAATGGGACACAGGAAATTGGGCCCCGTGGGTATGACTGAGA
ACTGCAATAGGAAAGATTTTCTGACCTGTCAATCGGGGGACCACCATTATGACACACGGAAAC
TTGGCTCAAGAACCCACTGATTGGACCACATCGAATACCGAGATGTGCGAGGTGGGGCAGGT
GTGTCAGGAGACGCTGCTGCTCATAGATGTAGGACTCACATCAACCCTGGTGGGGACAAAAG
GCTGCAGCACTGTTGGGGCTCAAAATTCCCAGAAGACCACCATCCACTCAGCCCCCTCCTGGG
GTGCTTGTGGCCTCCTATACCCACTTCTGCTCCTCGGACCTGTGCAATAGTGCCAGCAGCAG
CAGCGTTCTGCTGAACTCCCTCCCTCCTCAAGCTGCCCCCTGTCCCAGGAGACCGGCAGTGTC
CTACCTGTGTGACGCCCTTGGAACCTGTTCAAGTGGCTCCCCCGAATGACCTGCCCCAGG
GGCGCCACTCATTGTTATGATGGGTACATTCACTCTCTCAGGAGGTGGGCTGTCCACCAAAT
GAGCATTCAGGGCTGCGTGGCCCAACCTTCCAGCTTCTTGTTGAACCACACCAGACAAATCG
GGATCTTCTCTGCGCGTGAGAAGCGTGATGTGCAGCCTCCTGCCTCTCAGCATGAGGGAGGT
GGGGCTGAGGGCCTGGAGTCTCTCACTTGGGGGGTGGGGCTGGCACTGGCCCCAGCGCTGTG
GTGGGGAGTGTTTTGCCCTTCTGCTTAACTCTATTACCCCCACGATTCTTCACCGCTGCTGA
CCACCCACACTCAACCTCCCTCTGACCTCATAACCTAATGGCCTTGGACACCAGATTCTTTC
CCATTCTGTCCATGAATCATCTTCCCCACACACAATCATTCATATCTACTCACCTAACAGCA
AACTGGGGAGAGCCTGGAGCATCCGGAATTGCCCTATGGGAGAGGGGACGCTGGAGGAGTG
GCTGCATGTATCTGATAATACAGACCCTGTCCTTTCA

FIGURE 251

GCGACGGGCAGGACGCCCCGTTGCGCTAGCGCGTGCTCAGGAGTTGGTGTCTGCCTGCGCT
CAGG**ATG**AGGGGGAATCTGGCCCTGGTGGGCGTTCTAATCAGCCTGGCCTTCCTGTCACTGCTG
CCATCTGGACATCCTCAGCCGGCTGGCGATGACGCCTGCTCTGTGCAGATCCTCGTCCCTGG
CCTCAAAGGGGATGCGGGAGAGAAGGGAGACAAAGGCGCCCCCGACGGCCTGGAAGAGTCG
GCCCCACGGGAGAAAAAGGAGACATGGGGGACAAAGGACAGAAAGGCAGTGTGGGTGTCAT
GGAAAAATTGGTCCCATTGGCTCTAAAGGTGAGAAAGGAGATTCCGGTGACATAGGACCCCC
TGGTCTTAATGGAGAACCAGGCCTCCCATGTGAGTGACGCCAGCTGCGCAAGGCCATCGGGG
AGATGGACAACCAGGTCTCTCAGCTGACCAGCGAGCTCAAGTTCATCAAGAATGCTGTGCGC
GGTGTGCGCGAGACGGAGAGCAAGATCTACCTGCTGGTGAAGGAGGAGAAGCGCTACGCGGA
CGCCCAGCTGTCCTGCCAGGGCCGCGGGGGCACGCTGAGCATGCCCAAGGACGAGGCTGCCA
ATGGCCTGATGGCCGCATACCTGGCGCAAGCCGGCCTGGCCCGTGTCTTCATCGGCATCAAC
GACCTGGAGAAGGAGGGCGCCTTCGTGTACTCTGACCACTCCCCCATGCGGACCTTCAACAA
GTGGCGCAGCGGTGAGCCCAACAATGCCTACGACGAGGAGGACTGCGTGAGATGGTGGCCT
CGGGCGGCTGGAACGACGTGGCCTGCCACACCACCATGTACTTCATGTGTGAGTTTGACAAG
GAGAACATG**TGA**GCCTCAGGCTGGGGCTGCCCATTTGGGGGCCCCACATGTCCCTGCAGGGTT
GGCAGGGACAGAGCCCAGACCATGGTGCCAGCCAGGGAGCTGTCCCTCTGTGAAGGGTGGAG
GCTCACTGAGTAGAGGGCTGTTGTCTAAACTGAGAAAATGGCCTATGCTTAAGAGGAAAATG
AAAGTGTTCCCTGGGGTGCTGTCTCTGAAGAAGCAGAGTTTCATTACCTGTATTGTAGCCCCA
ATGTCATTATGTAATTATTACCCAGAATTGCTCTTCATAAAGCTTGTGCCTTTGTCCAAGC
TATACAATAAAATCTTTAAGTAGTGCAGTAGTTAAGTCCAAAAAAAAAAAAAAAAAAAAA

FIGURE 252

M R G N L A L V G V L I S L A F L S L L P S G H P Q P A G D D A C S V Q I L V P G L K G D A G E K G D K G A P G R P G R V G
P T G E K G D M G D K G Q K G S V G R H G K I G P I G S K G E K G D S G D I G P P G P N G E P G L P C E C S Q L R K A I G E
M D N Q V S Q L T S E L K F I K N A V A G V R E T E S K I Y L L V K E E K R Y A D A Q L S C Q G R G G T L S M P K D E A A N
G L M A A Y L A Q A G L A R V F I G I N D L E K E G A F V Y S D H S P M R T F N K W R S G E P N N A Y D E E D C V E M V A S
G G W N D V A C H T T M Y F M C E F D K E N M

FIGURE 253

AGTGACTGCAGCCTTCCTAGATCCCCCTCCACTCGGTTTCTCTCTTTGCAGGAGCACCGGCAG
CACCAGTGTGTGAGGGGAGCAGGCAGCGGTCCTAGCCAGTTCCTTGATCCTGCCAGACCACC
CAGCCCCCGGCACAGAGCTGCTCCACAGGCACC**ATG**AGGATCATGCTGCTATTACAGCCAT
CCTGGCCTTCAGCCTAGCTCAGAGCTTTGGGGCTGTCTGTAAGGAGCCACAGGAGGAGGTGG
TTCCTGGCGGGGGCCGCAGCAAGAGGGATCCAGATCTCTACCAGCTGCTCCAGAGACTCTTC
AAAAGCCACTCATCTCTGGAGGGATTGCTCAAAGCCCTGAGCCAGGCTAGCACAGATCCTAA
GGAATCAACATCTCCCGAGAAACGTGACATGCATGACTTCTTTGTGGGACTTATGGGCAAGA
GGAGCGTCCAGCCAGAGGGAAAGACAGGACCTTCTTACCTTCAGTGAGGGTTCCTCGGCCC
CTTCATCCCAATCAGCTTGGATCCACAGGAAAGTCTTCCCTGGGAACAGAGGAGCAGAGACC
TTTA**TAA**GACTCTCCTACGGATGTGAATCAAGAGAACGTCCCCAGCTTTGGCATCCTCAAGT
ATCCCCCGAGAGCAGAATAGGTACTCCACTTCCGGACTCCTGGACTGCATTAGGAAGACCTC
TTTCCCTGTCCCAATCCCCAGGTGCGCACGCTCCTGTTACCCTTTCTCTTCCCTGTTCTTGT
AACATTCTTGTGCTTTGACTCCTTCTCCATCTTTTCTACCTGACCCTGGTGTGGAAACTGCA
TAGTGAATATCCCCAACCCAATGGGCATTGACTGTAGAATACCCTAGAGTTCCTGTAGTGT
CCTACATTAAAAATATAATGTCTCTCTCTATTCCTCAACAATAAAGGATTTTTCATATGAA
AA

FIGURE 254

MRIMLLFTAILAFSLAQSFQAVCKEPQEEVVPGGGRSKRDPDLYQLLQRLFQSHSSLEGLLK
ALSQASTDPKESTSPEKRDHDFVGLMGKRSVQPEGKTGPFLPSVRVPRPLHPNQLGSTGK
SSLGTEEQRPL

Important features:

Signal peptide:

amino acids 1-18

Tyrosine kinase phosphorylation site.

amino acids 36-45

N-myristoylation site.

amino acids 33-39, 59-65

Amidation site.

amino acids 90-94

Leucine zipper pattern.

amino acids 43-65

Tachykinin family signature.

amino acids 86-92

FIGURE 255

GGGCGTCTCCGGCTGCTCCTATTGAGCTGTCTGCTCGCTGTGCCCCTGTGCCTGCTGTGCC
CGCGCTGTGCGCCGCTGCTACCGCGTCTGCTGGACGCGGGAGACGCCAGCGAGCTGGTGATTG
GAGCCCTGCGGAGAGCTCAAGCGCCCAGCTCTGCCCCAGGAGCCCAGGCTGCCCCGTGAGTC
CCATAGTTGCTGCAGGAGTGGAGCCATGAGCTGCGTCCTGGGTGGTGTTCATCCCCTTGGGGC
TGCTGTTCCCTGGTCTGCGGATCCCAAGGCTACCTCCTGCCCCAACGTCACTCTCTTAGAGGAG
CTGCTCAGCAAATACCAGCACAAACGAGTCTCACTCCCGGGTCCGCAGAGCCATCCCCAGGGA
GGACAAGGAGGAGATCCTCATGCTGCACAACAAGCTTCGGGGCCAGGTGCAGCCTCAGGCCT
CCAACATGGAGTACATGGTGAGCGCCGGCTCCGGCCGCAGAGGCTGGCACCGGGGGTGGGGC
CTGGGGCCACCAGCCTGCTCTGTTCCCCAGCCAGCTCTGTTCCCCAGCCAGTGCGTGTGATGG
CTGGCTCAGGGTCTCCTCTGGCAGGGGAGGATCCCGGCTCTGTTCTGTTTTGTTTGTGTTGTT
TTGAGACAGGGTCTCACTCTGCCACTGACGCTGGAGTGCAATGGCACAATCGTCATGCCCTG
AAACCTTAGACTCCCGGGGTAAAGCGATCCTGCTTCAGCCTCCCAAGTAGCTGGAACCTACAG
GCATGCACCATGGTGCCAGCTAGATTTTAAATATTTTGTGGAGATGGGGGTCTTGCTACGT
TGCCCAGGCTGGTCTTGAACTCCTAGGCTCAAGCAATCCTCCTGCCTCAGCCTCTCAAAGTG
CTAGGATTATAGGCATGAGTCACCCTGTCTGGCTCTGGCTCTGTTCTTAACATTCTGCCAAA
ACAACACACGTGGGTTCCTGTGCAGAGCCTGCCTCGTTGCCTTCATGTCACTCTTGGTAGC
TCCACTGGGAACACAGCTCTCAGCCTTTCCACCTGGAGGCAGAGTGGGGAGGGGGCCCAGGG
CTGGGCTTTGCTGATGCTGATCTCAGCTGTGCCACACGCTAGCTGCACCACCCTGACTTCTC
CTTAGCCCGTGTGAGCCTCACTTTCCACTTGAGAGTCCCTCCTCGCGTGGTTGCCATGACT
GTGAGATAAGTCGAGGCTGTGAAGGGCCCCGGCACAGACTGACCTGCCTCCCCAACCCCTAGG
CTTTGCTAACCGGGAAAGGAGCTAACGGTGACAGAAGACAGCCAAGGTCAACCCTCCCGGGT
GATTGTGATGGGTGTTCCAGGTGTGGTTGGGCGATGCTGCTACTTGACCCCAAGCTCCAGTG
TGGAACCTTCCTTCCTGGCTGGTTTTCCAGAACTACAGAGGAATGGACCACAGTCTTCCAGG
GTCCCTCCTCGTCCACCAACCGGGAGCCTCCACCTTGGCCATCCGTCAGCTATGAATGGCTT
TTTAAACAAACCCACGTCCCAGCCTGGGTAAACATGGTAAAGCCCCGTCTCTACAAAAAATC
CAAGTTAGCCGGGCATGGTGGTGCGCACCTGTAGTCCCAGCTGCAGTGGGACTGAGGTGGAG
GTGGAGGTGGGGGGTGGGAGCTGAGGAAGGAGGATCGCTTGAGCCTGGGAAGTCGAGGCTGC
AGTGAGCTGAGATTGCACCACTGCACTCCAGCCTGGGTGACAGAGCAAGACCCTGTCTCAAAAA

256/330

FIGURE 256

MSCVLGGVIPLEGLLFLVCGSQGYLLPNVTLLLEELLSKYQHNESHRSVRRRAIPREDKEEILML
HNKLRGQVQPQASNMEYMVSAGSGRRGWHRGWGLGHQPALFPSQLCSPASACDGWLRVSSGR
GGSRLCSVLFVCFETGSHSATDAGVQWHNRHALKP

Important features:

Signal peptide:

amino acids 1-22

N-glycosylation site.

amino acids 27-31, 41-45

N-myristoylation site.

amino acids 126-132, 140-146

Amidation site.

amino acids 85-89

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 257

AAGGAGAGGCCACCGGGACTTCAGTGTCTCCTCCATCCCAGGAGCGCAGTGGCCACTATGGG
GTCTGGGCTGCCCCCTTGTCTCCTCTTGACCCTCCTTGGCAGCTCACATGGAACAGGGCCGG
GTATGACTTTGCAACTGAAGCTGAAGGAGTCTTTTCTGACAAATTCCTCCTATGAGTCCAGC
TTCTTGGAATTGCTTGAAAAGCTCTGCCTCCTCCTCCATCTCCCTTCAGGGACCAGCGTCAC
CCTCCACCATGCAAGATCTCAACACCATGTTGTCTGCAACACATGACAGCCATTGAAGCCTG
TGTCTTCTTGGCCCGGGCTTTTGGGCCGGGGATGCAGGAGGCAGGCCCCGACCCTGTCTTT
CAGCAGGCCCCCACCCTCCTGAGTGGCAATAAATAAAATTTCGGTATGCTG

258/330

FIGURE 258

MSGGLPLVLLLTLLGSSHGTGPGMTLQLKLKESFLTNSSESSFLEELLEKLCLLLHLPSGTS
VTLHHARSQHHVVCNT

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000

FIGURE 259

AATTGTATCTGTGTAATGTTAAACAAACGAAATAAAATAGAAGGAAAACTTTCTGAGTTT
CAAAAACAACAGACTAGTACTCTAAAGAACTCTTTAAACAATTAAGTGTAGGATTGCAGT
TATGATTGGATATTATTTAATTCTGTTTCTGATGTGGGGTTCCTCCACTGTGTTCTGTGTGC
TATTAATATTTACCATTGCAGAAGCTTCATTCAGTGTTGAAAATGAATGCTTAGTGGATCTG
TGCCCTCTTACGCATATGTTACAAATTATCTGGAGTTCCTAATCAATGCAGAGTTCCCCCTCC
CTCCGATTGTTCTAAATAATTGAAAGATGTCTGCTGTGGAAAAAGGCATGTATTTAAATCTG
TATGATTCTCAACCATCTTTAGTTGGGAAAGGTCCTTGAAAGCCAATGGAAATACTTTTTTT
TTTTCTTGGCACTAATCAAGTGAGTGTTACCTTTTCACTTAGTAGGATGTGTTGTTACGCTA
GTAAAATAGAAACCTGTGTTTATTCTCAGGTATTTTAGAAACAACAGCCATCATTTTATTTT
ATGTGTGTGTTCTTGGCTGTATTCATAAATTATATATTTTGGGCTATCAAATATTACTTCAT
TCAATATAAATAACAATAGTAGAAGTTGTTTACTTAGATATGCTTTCTAGTTGCATTTTCTC
AGCCTATGTAAGACTACTTTGTTGTAATAGCCTTTGAAATTTACAGTACTGTCTCTCTACTA
TCTTCAGATTACTTGATTCAAATAAACCAATTATGTTTGTAATTGATATTAATAAAACCAGA
ATAAAAGTTCATATCTACCC

FIGURE 262

MMVALRGASALLVLFLAAFLPPPQCTQDPAMVHYIYQRFVLEQGLEKCTQATRAYIQEFQE
FSKNISVMLGRCQTYTSEYKSAVGNLALRVERAQREIDYIQYLREADECIVSEDKTLAEMLL
QEAEKKIRTLNASCDNMLMGIKSLKIVKKMMDTHGSWMKDAVYNPKVYLLIGSRNNTV
WEFANIRAFMEDNTKPAPRKQILTLWQGTGQVIYKGFLFFHNQATSNEIIKYNLQKRTVED
RMLLPGGVGRALVYQHSPSTYIDLAVDEHGLWAIHSGPGTHSHLVLTKEPGTLGVEHSWDT
PCRSQDAEASFLLCGVLYVVYSTGGQGPHRITCIYDPLGTISEEDLPNLFFPKRPRSHSMIH
YNPRDKQLYAWNEGNQIIYKLQTKRKLPLK

FIGURE 264

MELSQMSELMGLSVLLGLLALMATAAVARGWLRAGEERSGRPACQKANGFPPDKSSGSKKQK
QYQRIRKEKPQQHNFTHRLLAALKSHSGNISCMDFSSNGKYLATCADDRTIRIWSTKDFLQ
REHRSMRANVELDHATLVRFSPDCRAFIVWLANGDTLRVFKMTKREDGGYTFTATPEDFPKK
HKAPVIDIGIANTGKFIMTASSDTTVLIWSLKGQVLSTINTNQMNNTHA AVSPCGRFVASC
GFTPDVKVWEVCFGKKGEFQEVVRAFELKGHSA AVHSFAFSNDSRRMASVSKDGTWKLWDTDV
EYKKKQDPYLLKTGRFEEAAGAAPCRLALSPNAQVLALASGSSIHLYNTRRGEKEECFERVH
GECIANLSFDITGRFLASCGDRAVRLFHNTPGHRAMVEEMQGHLKRASNESTRQRLQQQLTQ
AQETLKSLGALKK

Important features:**Signal peptide:**

amino acids 1-25

N-glycosylation site.

amino acids 76-80, 92-96, 231-235, 289-293, 378-382, 421-425

Beta-transducin family Trp-Asp repeat protein.

amino acids 30-47, 105-118, 107-119, 203-216, 205-217, 296-308

266/330

FIGURE 266

MRGATRVSIMLLLVTVSDCAVITGACERDVQCGAGTCCAISLWLRGLRMCTPLGREGECHP
GSHKVPFFRKRKHHTCPCLPNLLCSRFPDGRYRCSMDLKNINF

Signal peptide:

amino acids 1-19

Tyrosine kinase phosphorylation site:

amino acids 88-95

N-myristoylation sites:

amino acids 33-39, 35-41, 46-52

266/330

FIGURE 268

MSFLQDPSFFTGMWSIGAGALGAAALALLLANTDVFLSKPQKALEYLEDIDLKTLEKEPR
TFKAKELWEKNGAVIMAVRRPGCFLCREEAADLSSLKSMLDQLGVPLYAVVKEHIRTEVKDF
QPYFKGEIFLDEKKKFYGPQRRKMMFMGFIRLGVWYNFFRAWNGGFSGNLEGEFGLGGVFV
VGSGKQGILLEHREKEFGDKVNLLSVLEAAKMIKPQTLASEKK

FIGURE 269

ACGGACCGAGGGTTTCGAGGGAGGGACACGGACCAGGAACCTGAGCTAGGTCAAAGACGCCCGG
GGCCAGGTGCCCCGTCGCAGGTGCCCTGGCCGGAGATGCGGTAGGAGGGGCGAGCGCGAGA
AGCCCCCTTCCTCGGCGCTGCCAACCCGCCACCCAGCCCATGGCGAACCCCGGGCTGGGGCTG
CTTCTGGCGCTGGGCCTGCCGTTCTTGCTGGCCCGCTGGGGCCGAGCCTGGGGGCAAATACA
GACCACTTCTGCAAATGAGAATAGCACTGTTTTGCCTTCATCCACCAGCTCCAGCTCCGATG
GCAACCTGCGTCCGGAAGCCATCACTGCTATCATCGTGGTCTTCTCCCTCTTGGCTGCCTTG
CTCCTGGCTGTGGGGCTGGCACTGTTGGTGCGGAAGCTTCGGGAGAAGCGGCAGACGGAGGG
CACCTACCGGCCAGTAGCGAGGAGCAGTTCTCCCATGCAGCCGAGGCCCGGGCCCCCTCAGG
ACTCCAAGGAGACGGTGCAGGGCTGCCTGCCCATCTAGGTCCCCCTCTCCTGCATCTGTCTCC
CTTCATTGCTGTGTGACCTTGGGGAAAGGCAGTGCCCTCTCTGGGCAGTCAGATCCACCCAG
TGCTTAATAGCAGGGAAGAAGGTACTTCAAAGACTCTGCCCCTGAGGTCAAGAGAGGATGGG
GCTATTCACTTTTATATATTTATATAAAATTAGTAGTGAGATGTAAAAAAAAAAAAAAAAAAAA

MANPGLGLLLALGLPFLRLRWGRAWGQIQTTSANENSTVLPSSSTSSSSDGNLRPEAITAIIV
VFSLLAALLLAVGLALLVRKLREKRQTEGTYRPSSEEQFSHAAEARAPQDSKETVQGCLPI

FIGURE 271

AATATATCATCTATTTATCATTAATCAATAATGTATTCTTTTATTCCAATAACATTTGGGTT
TTGGGATTTTAATTTTCAAACACAGCAGAAATGACATTTTTTCTGTCACTATTATTATTGTTG
GTATGTGAAGCTATTTGGAGATCCAATTCAGGAAGCAACACATTGGAGAATGGCTACTTTCT
ATCAAGAAATAAAGAGAACCACAGTCAACCCACACAATCATCTTTAGAAGACAGTGTGACTC
CTACCAAAGCTGTCAAAACCACAGGCAAGGGCATAGTTAAAGGACGGAATCTTGACTCAAGA
GGGTTAATTCTTGGTGCTGAAGCCTGGGGCAGGGGTGTAAAGAAAAACACTTAGATTCAATG
ATTGTAAATTTAAGGCAAATACACATATTAGTATTACCTTAGTGTAATGTATCCCTGTCATA
TATACAATAAGGTGAAATTATAAGTACCCTATGCAGTTGGCTGGACAGTTCTAAATTGGACT
TTATTAATTTTTTAAAATCAGTAACTGATTTATCACTGGCTATGTGCTTAGATCTACAGGAGA
TCATATAATTTGATACAAATAAAAGAAAAGTGTCTCTCCCCTTACAGAATTGACATTTTAA
ATGCGATACAGTTAGAATAGGAAATATGACATTAGAAAGGAAGAATGACAGGGAGAAAGGAA
AGAAGGGAAAATGTTGCCAAGGAAAAAAAAA

272/330

FIGURE 272

MTFFLSLLLLLVCEAIWRSNSGSNTLENGYFLSRNKENHSQPTQSSLEDSVTPTKAVKTTGK
GIVKGRNLDSRGLILGAEAWGRGVKKNT

[illegible]

GCCAGGAATAACTAGAGGAGGAACA**ATG**GGGTTATTTCAGAGGTTTTGTTCCTCTTAGTTCT
GTGCGCTGCTGCACAGTCAAATACTTCCCTTCATTAAGCTGAATAATAATGGCTTTGAAGATA
TTGTTCATTGTTATAGATCCTAGTGTGCCAGAAGATGAAAAAATAATTGAACAAATAGAGGAT
ATGGTGACTACAGCTTCTACGTACCTGTTTGAAGCCACAGAAAAAGATTTTTTTTTCAAAAA
TGTATCTATATTAATTCTGTGAGAATTGGAAGGAAAACTCTCAGTACAAAAGGCCAAAACATG
AAAACCATAAACATGTCTGATGTTATAGTTGCAACCACTACACTCCCAGGTAGAGATGAACCA
TACACCAAGCAGTTTACAGAATGTGGAGAGAAAGGCGAATACATTCACTTCACCCCTGACCT
TCTACTTGGAaaaaaaACAAATGAATATGGACCACCAGGCAAACCTGTTTGTCCATGAGTGGG
CTCACCTCCGGTGGGGAGTGTGTGATGAGTACAATGAAGATCAGCCTTTCTACCGTGTCTAAG
TCAAAAAAAATCGAAGCAACAAGGTGTTCCGCGAGGTATCTCTGGTAGAAAATAGAGTTTATAA
GTGTCAAGGAGGCAGCTGTCTTAGTAGAGCATGAGAATTGATTTCTACAACAAAACGTGATG
GAAAGATTTGTCAATTCTTTCTGTATAAAGTACAACACAGAAAAAGCATCCATAATGTTTATG
CAAAGTATTGATTCTGTGTGTTGAATTTTGTAAACGAAAAAACCCATAATCAAGAAGCTCCAAG
CCTACAAAACATAAAGTGAATTTTAGAAGTACATGGGAGGTGATTAGCAATTTCTGAGGATT
TTAAAAACACCATAACCCATGGTGACACCACCTCCTCCACCTGTCTTCTCATTTGCTGAAGATC
AGTCAAAGAATTGTGTGTTAGTTCTTGATAAGTCTGGAAGCATGGGGGTAAAGGACCGCCT
AAATCGAATGAATCAAGCAGCAAAAACATTTCTGCTGCAGACTGTTGAAAATGGATCCTGGG
TGGGGATGGTTCACTTTGATAGTACTGCCACTATTGTAAATAAGCTAATCCAATAAAAAGC
AGTGATGAAAGAAACACACTCATGGCAGGATTACCTACATATCCTCTGGGAGGAACTTCCAT
CTGCTCTGGAATTTAAATATGCATTTTCAGGTGATTGGAGAGCTACATTTCCCACTCGATGGAT
CCGAAGTACTGCTGCTGACTGATGGGGAGGATAAACACTGCAAGTTCTTGATTTGATGAAGTG
AAGACAAAGTGGGGCCATTGTTTATTGTTGTTTTGGGAAGAGCTGCTGATGAAGCAGTAAT
AGAGATGAGCAAGATAACAGGAGGAAGTCATTTTTATGTTTCAGATGAAGCTCAGAACAATG
GCCTCATTGATGCTTTTTGGGGCTCTTACATCAGGAAATACTGATCTCTCCAGAAAGTCCCTT
CAGCTCGAAAGTAAGGGATTAAACACTGAATAGTAATGCCTGGATGAACGACACTGTCATAAT
TGATAGTACAGTGGGAAGGACAGTTCTTTCTCATCATGGAACAGCTGCCTCCCGATA
TTTTCTCTCTGGGATCCAGTGGAACAATAATGGAATAATTCACAGTGGATGCAACTTCCAAA
ATGGCCCTATCTCAGTATTCAGGAACTGCAAAGGTGGGCACCTTGGGCATACAATCTTCAAGC
CAAAGCGAACCAGAAACATTAACATTTACAGTAACTTCTCGAGCAGCAAATTTCTTCTGTGC
CTCCAATCACAGTGAATGCTAAAATGAATAAGGACGTAAACAGTTTCCCCAGCCCCAATGATT
GTTTACGCAAAATTTCTACAAGGATATGTACCTGTTCTTGGAGCCAATGTACTGCTTTTAT
TGAATCACAGAATGGACATACAGAAGTTTTTGAACCTTTTGATAATGGTGCAGGCCTGATT
CTTTCAAGAATGATGGAGTCTACTCCAGGTATTTTACAGCATATACAGAAAATGGCAGATAT
AGCTTAAAAGTTTCGGGCTCATGGAGGAGCAAACACTGCCAGGCTAAAATTACGGCCTCCACT
GAATAGAGCCGCGTACATACCAGGCTGGGTAGTGAACGGGGAAATGGAAGCAAACCCGCCAA
GACCTGAAATTTGATGAGGATACTCAGACACCTTGGAGGATTTACGCCGAACAGCATCCGGA
GGTGCAATTTGTGGTATCACAAAGTCCCAAGCCTTCCCTTGCTGACCAATACCCCAAGTCA
AATCACAGACCTTGATGCCACAGTTCATGAGGATAAGATTATTTCTTACATGGACAGCACCAG
GAGATAATTTTGATGTTGAAAAGTTCAACGTTATATCATAAGAATAAGTGAAGTATTCTT
GATCTAAGAGCAGTTTTGATGATGCTCTTCAAGTAAATACTACTGATCTGTACCAAAAGGA
GGCCAACCTCAAGGAAAGCTTTGCATTTAAACAGAAAATATCTCAGAAGAAAATGCAACCC
ACATATTTATTGCCATTAAGAATATAGATAAAAGCAATTTGACATCAAAAGATATCCAACATT
GCACAAGTAACTTTGTTTTATCCCTCAAGCAAACTCTGATGACATTGATCCTACACCTACTCC
TACTCCTACTCCTACTCCTGATAAAAGTCATAATTTCTGGAGTTAATATTTCTACGCTGGTAT
TGTCTGTGATTGGGTCTGTTGTAATTGTTAACTTTATTTTTAAGTACCACCATT**TGA**ACCTTA
ACGAAGAAAAAAATCTTCAAGTAGACCTAGAAGAGAGTTTTAAAAAACAAAATGTAAGT
AAAGGATATTTCTGAATCTTAAATTCATCCCATGTGTATCATAAATCATATAAAATAATTT
TTAAGATGTCGAAAGGATACTTTGATTAAATAAAAAACACTCATGGATATGTAAAAACTGT
CAAGATTAAAATTTAATAGTTTTCATTTATTTGTTATTTTATTTGTAAGAAATAGTGATGAAC
AAAGATCCTTTTTTCATACTGATACCTGGTTGTATATTATTTGATGCAACAGTTTTCTGAAAT
GATATTTTCAAATTTGCATCAAGAAATTTAAATCATCTATCTGAGTAGTCAAATACAAGTAA
GGAGAGCAAAATAACAACATTTGGAAAAAaaaaaaAaaaaaaAaaaaaaAaaaaaaA
AaaaaaaAaaaaaaAaaaaaaAaaaaaaAaaaaaaAaaaaaaAaaaaaaAaaaaaaA

FIGURE 274

MGLFRGFVFLVLCLLHQSNSTSFIKLNNNGFEDIVIVIDPSVPEDEKIIEQIEDMVTASTY
LFEATEKRFFFFKNVSILIPENWKENPQYKRPKHENHKHADVIVAPPTLPGRDEPYTKQFTEC
GEKGEYIHFTPDLLLGKKQNEYGPPGKLFVHEWAHLRWGVFDEYNEDQPFYRAKSKKIEATR
CSAGISGRNRVYKCQGGSCLSRACRIDSTTKLYGKDCQFFPDKVQTEKASIMFMQSIDSVE
FCNEKTHNQEAPSLQNIKCNRSTWEVISNSEDFKNTIPMVTPPPPPVFSLLKISQRIVCLV
LDKSGSMGGKDRLNRMNQAACHFLLQTVENGSWVGMVHFDSTATIVNKLIQIKSSDERNTLM
AGLPTYPLGGTSICSGIKYAFQVIGELHSQLDGSEVLLLTGDGEDNTASSCIDEVKQSGAIVH
FIALGRAADEAVIEMSKITGGSHFYVSDEAQNNGLIDAFGALTSGNTDLSQKSLQLESKGLT
LNSNAWMNDTVIIDSTVGKDTFFLITWNSLPPSISLWDPSGTIMENFTVDATSKMAYLSIPG
TAKVGTWAYNLQAKANPETLTITVTSRAANSSVPPITVNAKMNKDVNSFPSPMIVYAEILQG
YVPVLGANVTAFIESQNGHTEVLELLDNGAGADSFKNDBGVYSRYFTAYTENGRYSLKVRAGH
GANTARLKLRPPLNRAAYIPGWVNGEIEANPPRPEIDEDTQTTLEDFSRITASGGAFVVSQV
PSLPLPDQYPPSQITDLDATVHEDKIILTWTAPGDNFVDVGKVQRYIIRISASILDLRDSFDD
ALQVNTTDLSPKEANSKESFAFKPENISEENATHIFIAIKSIDKSNLTSKVSANIAQVTLFIP
QANPDDIDPTPTPTPTPTPDKSHNSGVNISTLVLSVIGSVVIVNFILSTTI

Signal peptide:

amino acids 1-21

Putative transmembrane domains:

amino acids 284-300, 617-633

Leucine zipper pattern.

amino acids 469-491, 476-498

N-glycosylation site.amino acids 20-24, 75-79, 340-344, 504-508, 542-546, 588-592,
628-632, 811-815, 832-836, 837-841, 852-856, 896-900

CTCCTTAGGTTGGAACCCCTGGGAGTAGAGTACTGACAGCAAAGACCGGGAAGACCATACGTCCTCCCG
GGCAGGGGTGACAACAGGTGTCATCTTTTTGATCTCGTGTGTGGCTGCCTTCTATTTCAAGGAAAG
ACGGCAAGGTAATTTTGACCCAGGAGGAATGATGTAGCCACCTCTCAACCTTCCCTTCTTGAACCT
CCGAGTTATGCCAGGATTTACTAGAGAGTGTCAACTCAACCCAGCAAGCGGCTCCTTCGGCTTAACCT
GTGGTTGGAGGAGAGAACCTTTGTGGGGCTGCGTTCTCTTAGCAGTGCTCAGAAGTGACTTGCCTGA
GGGTGGACCAGAAAGAAAGGTAAGGTTCCCTCTTGCTGTTGGCTGCACATCAGGAAGGCTGTGATGGG
AATGAAGGTGAAAACCTTTGGAGATTTCACTTCAGTCATTGCTTCTGCTGCAAGATCATCCTTTAAAA
GTAGAGAAGCTGCTCTGTGTGGTGGTTAACTCCAAGAGGCAGAAGCTGCTTCTAGAAGGAAATGGATG
CAAGCAGCTCCGGGGGCCCCAAACGCATGCTTCTGTGGTCTAGCCCAGGGAAGCCCTTCCGTGGGG
GCCCCGGCTTTGAGGGATGCCACCGGTTCTGGACGCATGGCTGATTCCTGAATGATGATGGTTCTGCC
GGGGGCTGCTTGCCTGGATTTCCCGGGTGGTGGTTTGTGCTGGTCTCTCTGCTGTGCTATCTCTGT
CCTGTACATGTTGGCCTGCACCCAAAAGGTGACGAGGAGCAGCTGGCACTGCCAGGGGCCAACAGC
CCCACGGGGAAGGAGGGGTACCAGGCCGTCTTCAGGAGTGGGAGGAGCAGCACCCGCAACTACGTGA
GCAGCCTGAAGCGGCAGATCGCACAGCTCAAGGAGGAGCTGCAGGAGAGGAGTGAGCAGCTCAGGAA
TGGGCAGTACCAAGCCAGCGATGCTGCTGGCTGGGTCTGGACAGGAGCCCCCAGAGAAAACCCAG
GCCAGCTCTCTGGCTTCTGCATCTCGCAGGTGGACAAGGCAGAGTGTAATGCTGGCGTCAAGCTGG
CCACAGAGTATGCAGCAGTGCTTTCGATAGCTTTACTCTACAGAAGGTGTACCAGCTGGAGACTGG
CCTTACCCGCCACCCCGAGGAGAAGCCTGTGAGGAAGGACAAGCGGGATGAGTTGGTGGGAAGCCATT
GAATCAGCCTTGGAGACCTTGAACAACCTCGACAGAGAACAGCCCAATCACCCTCTTACACGGCCT
CTGATTTTCATAGAAGGATCTACCGAACAGAAAGGGACAAGGGACATTGTATGAGCTCACCTTCAA
AGGGGACCACAAACACGAATTCAAACGGCTCATCTTATTTGACCATTACGCCCATCATGAAAGTG
AAAAATGAAAAGCTCAACATGGCCAACACGCTTATCAATGTTATCGTGCCTTAGCAAAAAGGGTGG
ACAAGTTCCGGCAGTTTCATGCAGAATTCAGGGAGATGTGCATCTGACAGGATGGGAGAGTCCATCT
CACTGTTGTTTACTTTTGGGAAAGAAGAAATAAATGAAGTCAAAGGAATACTTGAAAACACTTCCAAA
GCTGCCAACTTCAGGAACCTTACCTTCATCCAGCTGAATGGAGAATTTTCTCGGGGAAAGGGACTTGA
ATGTTGGAGCCCGCTTCTGGAAGGGAAGCAACGTCCTTCTCTTTCTGTGATGTGGACATCTACTT
CACATCTGAATTCCTCAATACGTGTAGGCTGAATACACAGCCAGGGAAGGATTTTATCCAGTT
CTTTTCAGTCAGTACAATCCTGGCATAATATACGGCCACCATGATGCAGTCCCTCCCTTGGAAACAG
AGCTGGTCATAAAGAAGGAAACTGGATTTTGGAGAGACTTTGGATTTGGGATGACGTGTGAGTATCG
GTCAGACTTCATCAATATAGTGGGTTTGTATCTGGACATCAAAGCTGGGGCGGAGGAGATGTGCAC
CTTTATCGCAAGTATCTCCACAGCAACCTCATAGTGGTACGACGCGCTGTGCGAGGACTCTCCACC
TCTGGCATGAGAAGCGCTGCATGGACGAGCTGACCCCGAGCAGTACAAGATGTGCATGCAGTCCAA
GGCCATGAACGAGGCATCCCACGGCCAGCTGGGCATGTGTTGTTTCAGGCACGAGATAGAGGCTCAC
CTTCGCAACACAGAAACAGAGAACAAGTAGCAAAAAAACATGAATCCCCAGAGAAGGATTTGGGAGA
CACTTTTTCTTTCTTTTGGCAATTACTGAAAGTGGCTGCAACAGAGAAAGACTTCCATAAAGGACG
ACAAAAGAATTGGACTGATGGGTGAGAGATGAGAAAGCCTCCGATTTCTCTCTGTTGGGCTTTTTAC
AACAGAAATCAAATCTCCGCTTTGCTGCAAAAGTAACCCAGTTGCAACCTGTGAAGTGTCTGACA
AAGGCAGAATGCTTGTGAGATTATAAGCCTAATGGTGTGGAGGTTTGTGATGTTTACAATACAT
GAGACCTGTGTTTTGTGTGCTCATTGAAATATTCTGATTTAAGAGCAGTTTGTAAAAAATTATCAT
TAGCATGAAAGGCAAGCATATTTCTCCTCATATGAATGAGCCTATCAGCAGGGCTCTAGTTTCTAG
AATGCTAAAATATCAGAAGGCAGGAGAGGAGATAGGCTTATATGATACTAGTGAGTACATTAAGTA
AAAAATAATGGACCAGAAAAGAAAAGAACCATAAATATCGTGCATATTTTCCCAAGATTAAACCA
AAAAATACTGCTTATCTTTTGGTGTGCTTTTAACTGTCTCCGTTTTTTTTCTTTTATTTAAAAAT
GCACTTTTTTTTCCCTTGTGAGTTATAGTCTGCTTATTTAATTACCATTGCAAGCTTTACAAGAGA
GCACAAGTTGGCCTACATTTTATATTTTAAAGAAGATACTTTGAGATGATGAGACTTTCA
GTTCAAGCATCAAATTGATGCCATATCCAAGGACATGCCAAATGCTGATTTCTGTGAGGCACTGAAT
GTCAGGCATTGAGACATAGGGAAGGAATGGTTTGTACTAATACAGACGTACAGATACTTTCTCTGAA
GAGTATTTTCAAGAGGAGCAACTGAACACTGGAGGAAAAGAAAATGACACTTTCTGTTTTACAGAA
AAGGAACTCATTAGACTGGTGATATCGTGATCTACCTAAAAGTCAGAAACCACATTTTCTCCTCA
GAAGTAGGGACCGCTTTCTTACCTGTTTAAATAAACCAAGATATACCGTGTGAACCAACAATCTCT
TTTTCAAAACAGGGTGTCTCCTCCTGGCTTCTGGCTTCCATAAGAAGAAATGGAGAAAATATATATAT
ATATATATATATTTGTAAAGATCAATCCATCTGCCAGAATCTAGTGGGATGGAAGTTTTTGTCTACAT
GTTATCCACCCAGCCAGGTGGAAGTAAGTCAATTTTAAATTAAGCAGTTCTACTCAATCA
CCAAGATGCTTCTGAAAATTGCAATTTTATACCATTTCAAACTATTTTTTAAAAATAAATACAGTTA
ACATAGAGTGGTTTCTTCATTTCATGTGAAAAATTATTAGCCAGCACCAGATGCATGAGCTAATTATCT
CTTTGAGTCTGTTGTTTTCTGTTGTGCTACAGTAACTCATTTGTTTAAAGCTTCAAGAACATTCAGC
TGTTGGTGTGTTAAAAAATGCATTGTATTGATTGTACTGGTAGTTTATGAAATTTAATTAACACAC
AGGCCATGAATGGAAGGTGGTATTGACACAGTAAATAAATATGATTTGTGGATATGAA

FIGURE 277

GAAAGAATGTTGTGGCTGCTCTTTTTCTGGTGACTGCCATTCATGCTGAACTCTGTCAACC
AGGTGCAGAAAATGCTTTTAAAGTGAGACTTAGTATCAGAACAGCTCTGGGAGATAAAGCAT
ATGCCTGGGATACCAATGAAGAATACCTCTTCAAAGCGATGGTAGCTTTCTCCATGAGAAAA
GTTCCCAACAGAGAAGCAACAGAAATTTCCCATGTCCTACTTTGCAATGTAACCCAGAGGGT
ATCATTCTGGTTTGTGGTTACAGACCCTTCAAAAAATCACACCCTTCCTGCTGTTGAGGTGC
AATCAGCCATAAGAATGAACAAGAACCGGATCAACAATGCCTTCTTTCTAAATGACCAAAC
CTGGAATTTTTTAAAAATCCCTTCACACTTGCACCACCCATGGACCCATCTGTGCCCATCTG
GATTATTATATTTGGTGTGATATTTTGCATCATCATAGTTGCAATTGCACTACTGATTTTAT
CAGGGATCTGGCAACGTAGAAGAAAGAACAAAGAACCATCTGAAGTGGATGACGCTGAAGAT
AAGTGTGAAAACATGATCACAATTGAAAATGGCATCCCCTCTGATCCCCTGGACATGAAGGG
GGGCATATTAATGATGCCTTCATGACAGAGGATGAGAGGCTCACCCCTCTCTGAAGGGCTGT
TGTTCTGCTTCCTCAAGAAATTAAACATTTGTTTCTGTGTGACTGCTGAGCATCCTGAAATA
CCAAGAGCAGATCATATATTTTGTTTCACCATTCTTCTTTTGAATAAATTTTGAATGTGCT
TGAAAGTGAAAAGCAATCAATTATACCCACCAACACCACTGAAATCATAAGCTATTCACGAC
TCAAAATATTCTAAATATTTTTCTGACAGTATAGTGTATAAATGTGGTCATGTGGTATTTG
TAGTTATTGATTTAAGCATTTTGTAGAAATAAGATCAGGCATATGTATATATTTTCACACTTC
AAAGACCTAAGGAAAAATAAATTTTCCAGTGGAGAATACATATAATATGGTGTAGAAATCAT
TGAAAATGGATCCTTTTTGACGATCACTTATATCACTCTGTATATGACTAAGTAAACAAAAG
TGAGAAGTAATTATTGTAAATGGATGGATAAAAATGGAATTACTCATATACAGGGTGGAATT
TTATCCTGTTATCACACCAACAGTTGATTATATATTTTCTGAATATCAGCCCCCTAATAGGAC
AATTCTATTTGTTGACCATTTCTACAATTTGTAAAAGTCCAATCTGTGCTAACTTAATAAAG
TAATAATCATCTCTTTTTTAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 279

AACTCAAACCTCCTCTCTCTGGGAAAACGCGGTGCTTGCTCCTCCCGGAGTGGCCTTGGCAGG
GTGTTGGAGCCCTCGGTCTGCCCCGTCCGGTCTCTGGGGCCAAGGCTGGGTTTCCCTC**ATGT**
ATGGCAAGAGCTCTACTCGTGCGGTGCTTCTTCTCCTTGGCATAACAGCTCACAGCTCTTTGG
CCTATAGCAGCTGTGGAAATTTATACCTCCCGGGTGCTGGAGGCTGTTAATGGGACAGATGC
TCGGTTAAAATGCACTTTCTCCAGCTTTGCCCCCTGTGGGTGATGCTCTAACAGTGACCTGGA
ATTTTCGTCCTCTAGACGGGGGACCTGAGCAGTTTGTATTCTACTACCACATAGATCCCTTC
CAACCCATGAGTGGGCGGTTTAAGGACCGGGTGTCTTGGGATGGGAATCCTGAGCGGTACGA
TGCTCCATCCTTCTCTGGAACTGCAGTTCGACGACAATGGGACATACACCTGCCAGGTGA
AGAACCACCTGATGTTGATGGGGTGATAGGGGAGATCCGGCTCAGCGTCGTGCACACTGTA
CGCTTCTCTGAGATCCACTTCCTGGCTCTGGCCATTGGCTCTGCCTGTGCACTGATGATCAT
AATAGTAATTGTAGTGGTCCTCTTCCAGCATTACCGGAAAAAGCGATGGGCCGAAAGAGCTC
ATAAAGTGGTGGAGATAAAATCAAAGAAGAGGAAAGGCTCAACCAAGAGAAAAAGGTCTCT
GTTTATTTAGAAGACACAGAC**TAA**CAATTTTAGATGGAAGCTGAGATGATTTCCAAGAACAA
GAACCCTAGTATTTCTTGAAGTTAATGGAACTTTTCTTTGGCTTTTCCAGTTGTGACCCGT
TTTCCAACCAGTTCTGCAGCATATTAGATTCTAGACAAGCAACACCCCTCTGGAGCCAGCAC
AGTGCTCCTCCATATCACCAGTCATACACAGCCTCATTATTAAGGTCTTATTTAATTTCAGA
GTGTAAATTTTTTCAAGTGCTCATTAGGTTTTATAAACAAGAAGCTACATTTTTGCCCTTAA
GACACTACTTACAGTGTTATGACTTGTATACACATATATTGGTATCAAAGGGGATAAAAGCC
AATTTGTCTGTTACATTTCTTTTACGTATTTCTTTTAGCAGCACTTCTGCTACTAAAGTTA
ATGTGTTTACTCTCTTTCCTTCCCACATTCTCAATTAAAAGGTGAGCTAAGCCTCCTCGGTG
TTTCTGATTAACAGTAAATCCTAAATTCAACTGTAAATGACATTTTTATTTTTATGTCTC
TCCTTAACCTATGAGACACATCTTGTTTTACTGAATTTCTTTCAATATTCCAGGTGATAGATT
TTTGTCTG

280/330

FIGURE 280

MYGKSSTRAVLLLLGIQLTALWPAAVEIYTSRVLEAVNGTDARLKCTFSSEFAPVGDALTVT
WNFRPLDGGPEQFVFYYHIDPFQPMSGRFKDRVSWDGNPERYDASILLWKLQFDDNGTYTCQ
VKNPPDVGVI GEIRLSVVHTVRFSEIHFLALAIGSACALMIIIVIVVVL FQHYRKKRWAER
AHKVVEIKSKEERLNQEKKVSVYLEDTD

280/330

282/330

FIGURE 282

MKFLAVLVLLGVSI FLVSAQNPTTAAPADTYPATGPADDEAPDAETTAAATTATTAAPTAT
TAASTTARKDIPVLPKWVGDL PNGRVCP

090374 22/03/50

FIGURE 283

GGACTCTGAAGGTCCCAAGCAGCTGCTGAGGCCCCCAAGGAAGTGGTTCCAACCTTGGACCC
CTAGGGGTCTGGATTTGCTGGTTAACAAGATAACCTGAGGGCAGGACCCCATAGGGGA**ATGC**
TACCTCCTGCCCTTCCACCTGCCCTGGTGTTCACGGTGGCCTGGTCCCTCCTTGCCGAGAGA
GTGTCCTGGGTGAGGACGCAGAGGACGCTCACAGACTCCAGCCCTTTGTTACCGAGAGGAC
ACTTGGCAAGGTCCAGCGATGGTCCGGAGTCCACACACAGACTGGCGGCAGGGCAGGAGGGG
GACAGTTCTGTTGTGCTTGGTTGGACAGTAAGAGGGTCTTGGCCAGTCCAGGGTGGGGGGCG
GCAAACCTCCATAAAGAACCAGAGGGTCTGGGCCCCGGCCACAGAGTCATCTGCCCAGCTCCT
CTGCTGCTGGCCAGTGGGAGTGGCACGAGGTGGGGCTTTGTGCCAG**TAA**AACCACAGGCTGG
ATTTGCCTGCGGGCCATGGTCCCTGTCTAGGGCAGCAATTCTCAACCTTCTTGCTCTCAGGA
CCCCAAAGAGCTTTCATTGTATCTATTGATTTTACCACATTAGCAATTAAACTGAGAAAT
GGGCCGGGCACGGTGGCTCACGCCTGTAATCCCAGCACTTTGGGAGGCCGAGGCGGGTGGAT
CACCTGAGATCAGGAGTTCAAGACCAGCCTGGCCAACATGGTGAAACCTTGTCTACTAAAAA
TACAAAAAATTAGCCAGGCACAGTGGTGTGCACTGGTAGTCCCAGTTACTCGGGAGGCTGAG
GCAGGAAAATCGCTTGAACCCAGGAGGCGGACGTTGCGGTGAGCCGAGATCGCGCCGCTGAT
TCCAGCCTGGGCGACAAGAGTGAGACTCCATCTCACACA

FIGURE 285

GTC**ATG**CCAGTGCCTGCTCTGTGCCTGCTCTGGGCCCTGGCAATGGTGACCCGGCCTGCCTCA
GCGGCCCCCATGGGCGGCCAGAACTGGCACAGCATGAGGAGCTGACCCTGCTCTTCCATGG
GACCCTGCAGCTGGGCCAGGCCCTCAACGGTGTGTACAGGACCACGGAGGGACGGCTGACAA
AGGCCAGGAACAGCCTGGGTCTCTATGGCCGCACAATAGAACTCCTGGGGCAGGAGGTCAGC
CGGGGCCGGGATGCAGCCCAGGAACCTTCGGGCAAGCCTGTTGGAGACTCAGATGGAGGAGGA
TATTCTGCAGCTGCAGGCAGAGGCCACAGCTGAGGTGCTGGGGGAGGTGGCCCAGGCACAGA
AGGTGCTACGGGACAGCGTGCAGCGGCTAGAAGTCCAGCTGAGGAGCGCCTGGCTGGGCCCT
GCCTACCGAGAATTTGAGGTCTTAAAGGCTCACGCTGACAAGCAGAGCCACATCCTATGGGC
CCTCACAGGCCACGTGCAGCGGCAGAGGCGGGAGATGGTGGCACAGCAGCATCGGCTGCGAC
AGATCCAGGAGAGACTCCACACAGCGGCGCTCCCAGCC**TGA**ATCTGCCTGGATGGAAGTGA
GACCAATCATGCTGCAAGGAACACTTCCACGCCCCGTGAGGCCCTGTGCAGGGAGGAGCTG
CCTGTTCACTGGGATCAGCCAGGGCGCCGGGCCCCACTTCTGAGCACAGAGCAGAGACAGAC
GCAGGCGGGGACAAAGGCAGAGGATGTAGCCCCATTGGGGAGGGGTGGAGGAAGGACATGTA
CCCTTTCATGCCTACACACCCCTCATTAAGCAGAGTCGTGGCATTTCAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 287

GGCAAC**ATG**GCTCAGCAGGCTTGCCCCAGAGCCATGGCAAAGAATGGACTTGTAATTTGCAT
 CCTGGTGATCACCTTACTCCTGGACCAGACCACCAGCCACACATCCAGATTAAAAGCCAGGA
 AGCACAGCAAACGTCGAGTGAGAGACAAGGATGGAGATCTGAAGACTCAAATTGAAAAGCTC
 TGGACAGAAGTCAATGCCTTGAAGGAAATTCAAGCCCTGCAGACAGTCTGTCTCCGAGGCAC
 TAAAGTTCACAAGAAATGCTACCTTGCTTCAGAAGGTTTGAAGCATTTCATGAGGCCAATG
 AAGACTGCATTTCCAAAGGAGGAATCCTGGTTATCCCCAGGAACTCCGACGAAATCAACGCC
 CTCCAAGACTATGGTAAAAGGAGCCTGCCAGGTGTCAATGACTTTTGGCTGGGCATCAATGA
 CATGGTCACGGAAGGCAAGTTTGTGACGTCAACGGAATCGCTATCTCCTTCCTCAACTGGG
 ACCGTGCACAGCCTAACGGTGGCAAGCGAGAAAACGTGTGTCTTCTCCCAATCAGCTCAG
 GGCAAGTGGAGTGATGAGGCCTGTGCGCAGCAGCAAGAGATACATATGCGAGTTCACCATCCC
 TAAAT**TAG**GTCTTTCTCCAATGTGTCTCCAAGCAAGATTCATCATAACTTATAGGTTTCATGA
 TCTCTAAGATCAAGTAAAAATCATAATTTTTACTTATTAAAAAATTGCAACACAAGATCAAT
 GTCCATAGCAATATGATAGCATCAGCCAATTTTGCTAACACATTTCTTTGGGATTTTGCCCT
 TCCTGGGGTATAGGGGATCAGAAATATTGATCCATGTGCACGCAGATAAAATGGCTTCTGCT
 AAACAGACTAAAATCTTTCTCTAGTCTTTCTCACTTGTACAAACCCAGTTTGTTTTCAA
 AAATCACAGTAGCAATGCAACTCATCACTCTAGAAAAGCAAGCTTAGGCTACCTGAAAGATT
 TTCCCTTGGAAGTTTAGCGTATGTTTGACTAACAAAAATTCCTACATCAGAGACTCTAGGT
 GCTATATAATCCAAAACTTTTCAGCCTGTTGCTCATTCTGTCCCATGCTGGCAATAATACC
 TTGTCAGCCCATTACCCTTATTTTGAATTGCTCCATCTCCTGGTGGGACTTGTATCTTGTCT
 GCCATATCAGAACACAAACCCCTGAAGAGGTTCTGATTTGATTTTTTTTTTTTCTTCATGCC
 TACCCTTTTTTTTGGAAAGTTTCCAGCCGCAATTTGAAATGAAATGACAAGGTGTATATTTGAT
 CAATTTTCATTCCCACCATTGCATTACAACCTCTAACTTAAATGGGTAACCCTAAGGCATAT
 CAAAGAAGCAGATTGCATGATAAACGGAAATAGAAAAAAGAACCTACATTTATTTTGCTTT
 AGCATCCTTACTCTCACCTTTTATGAGATTGAGAGTGGACTTACATTTCTTTTTTACATTT
 TCGTATATTTATTTTTTTTAGCCATCATTATATGTTTAAGTCTATTATGGGCAACCAATCTT
 TGGAAGCTGAAAACCTGAATTTAAAGAATGCTATCTTGAAAATTGCATACGTCTGTGCAATT
 TTTTATTCTGCCTAGTGCTATTCTGCTTGTTTAACTAGATTGTACAAAATAACTTCATTGCT
 TAATATCAAATTACAAAGTTTAGACTTGGAGGGAAATGGGCTTTTGTAGAAGCAAACAATTTT
 AAATATATTTTGTCTTCAAATAAATAGTGTTTAAACATTGAATGTGTTTTGTGAACAATAT
 CCCACTTTGCAAACCTTAACTACACATGCTTGGAATTAAGTTTTAGCTGTTTTTCATTGCTCA
 ATAATAAAGCCTGAATTCTGATCAATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

288/330

FIGURE 288

MAQQACPRAMAKNGLVICILVITLLLDQTTSHTSRLKARKHKKRRVRDKDGD LKTQIEKLWT
EVNALKEIQALQTVCLRGTKVHKKCYLASEGLKHFHEANEDCISKGGILVIPRNSDEINALQ
DYGKRSLPGVNDFWLGINDMVTEGKFVDVNGIAISFLNWDRAQPNGGKRENCVLFSQSAQ GK
WSDEACRSSKRYICEFTIPK

288/330

FIGURE 289

GCGAGGACCGGGTATAAGAAGCCTCGTGGCCTTGCCCGGGCAGCCGCAGGTTCCCCGCGCGC
CCCCGAGCCCCCGCGCC**ATGA**AGCTCGCCGCCCTCCTGGGGCTCTGCGTGGCCCTGTCCTGCA
GCTCCGCTGCTGCTTTCTTAGTGGGCTCGGCCAAGCCTGTGGCCCAGCCTGTCGCTGCGCTG
GAGTCGGCGGGCGGAGGCCGGGGCCGGGACCCTGGCCAACCCCTCGGCACCCTCAACCCGCT
GAAGCTCCTGCTGAGCAGCCTGGGCATCCCCGTGAACCACCTCATAGAGGGCTCCCAGAAGT
GTGTGGCTGAGCTGGGTCCCCAGGCCGTGGGGGCCGTGAAGGCCCTGAAGGCCCTGCTGGGG
GCCCTGACAGTGTTTGGC**TGA**GCCGAGACTGGAGCATCTACACCTGAGGACAAGACGCTGCC
CACCCGCGAGGGCTGAAAACCCCGCCGCGGGGAGGACCGTCCATCCCCTTCCCCGGCCCCCT
CTCAATAAACGTGGTTAAGAGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAA

FIGURE 291

TGAAGGACTTTTCCAGGACCCAAGGCCACACACTGGAAGTCTTGCAGCTGAAGGGAGGCACT
 CCTTGGCCTCCGCAGCCGATCAC**ATGA**AGGTGGTGCCAAGTCTCCTGCTCTCCGTCTCTCCTG
 GCACAGGTGTGGCTGGTACCCGGCTTGGCCCCCAGTCCCTCAGTCGCCAGAGACCCCAGCCCC
 TCAGAACCAGACCAGCAGGGTAGTGCAGGCTCCCAGGGAGGAAGAGGAAGATGAGCAGGAGG
 CCAGCGAGGAGAAGGCCGGTGAGGAAGAGAAAAGCCTGGCTGATGGCCAGCAGGCAGCAGCTT
 GCCAAGGAGACTTCAAACCTTCGGATTTCAGCCTGCTGCGAAAGATCTCCATGAGGCACGATGG
 CAACATGGTCTTCTCTCCATTTGGCATGTCTTGGCCATGACAGGCTTGATGCTGGGGGCCA
 CAGGGCCGACTGAAACCCAGATCAAGAGAGGGCTCCACTTGCAGGCCCTGAAGCCCACCAAG
 CCCGGGCTCCTGCCTTCCCTCTTTAAGGGACTCAGAGAGACCTCTCCCGCAACCTGGAACCT
 GGGCCTCTCACAGGGGAGTTTTGCCTTCATCCACAAGGATTTTGATGTCAAAGAGACTTTCT
 TCAATTTATCCAAGAGGTATTTTGATACAGAGTGCCTGCCTATGAATTTTCGCAATGCCTCA
 CAGGCCAAAAGGCTCATGAATCATTACATTAACAAAGAGACTCGGGGGAAAATCCCAAACCT
 GTTTGATGAGATTAATCCTGAAACCAAATTAATTCCTTGTGGATTACATCTTGTTCAAAGGGA
 AATGGTTGACCCCATTTGACCCGTGTCTTCACCGAAGTCGACACTTTCCACCTGGACAAGTAC
 AAGACCATTAAAGGTGCCCATGATGTACGGTGCAGGCAAGTTTGCCTCCACCTTTGACAAGAA
 TTTTCGTTGTTCATGTCCTCAAACCTGCCCTACCAAGGAAATGCCACCATGCTGGTGGTCCTCA
 TGGAGAAAATGGGTGACCACCTCGCCCTTGAAGACTACCTGACCACAGACTTGGTGGAGACA
 TGGCTCAGAAACATGAAAACCAGAAACATGGAAGTTTTCTTTCCGAAGTTCAAGCTAGATCA
 GAAGTATGAGATGCATGAGCTGCTTAGGCAGATGGGAATCAGAAGAATCTTCTCACCCCTTG
 CTGACCTTAGTGAACCTCTCAGCTACTGGAAGAAATCTCCAAGTATCCAGGGTTTTACGAAGA
 ACAGTGATTGAAGTTGATGAAAGGGGGCACTGAGGCAGTGGCAGGAATCTTGTGAGAAATTAC
 TGCTTATTCCATGCCTCCTGTCATCAAAGTGGACCGGCCATTTCAATTCATGATCTATGAAG
 AAACCTCTGGAATGCTTCTGTTTCTGGGCAGGGTGGTGAATCCGACTCTCCTA**TAA**ATTCAGG
 ACATGCATAAGCACTTCGTGCTGTAGTAGATGCTGAATCTGAGGTATCAAACACACACAGGA
 TACCAGCAATGGATGGCAGGGGAGAGTGTTCCCTTTTGTTCCTTAAGTATAGGGTGTCTC
 AAATAAATACAGTAGTCCCCACTTATCTGAGGGGGATACATTCAAAGACCCCCAGCAGATGC
 CTGAAACGGTGGACAGTGCTGAACCTTATATATATTTTTTCTACACATACATACCTATGAT
 AAAGTTTAATTTATAAATTAGGCACAGTAAGAGATTAACAATAATAACAACATTAAGTAAAA
 TGAGTTACTTGAACGCAAGCACTGCAATACCATAACAGTCAAACCTGATTATAGAGAAGGCTA
 CTAAGTGAATCATGGGCGAGGAGCATAGACAGTGTGGAGACATTGGGCAAGGGGAGAATTCA
 CATCCTGGGTGGGACAGAGCAGGACGATGCAAGATTCCATCCCCTACTCAGAATGGCATGC
 TGCTTAAGACTTTTAGATTGTTTATTTCTGGAATTTTTTCATTTAATGTTTTTGGACCATGGT
 TGACCATGGTTAACTGAGACTGCAGAAAGCAAAACCATGGATAAGGGAGGACTACTACAAA
 GCATTAAATTGATACATATTTTTTAAAAA

FIGURE 292

MKVVP S L L L S V L L A Q V W L V P G L A P S P Q S P E T P A P Q N Q T S R V V Q A P R E E E E D E Q E A S E E K A G E
E E K A W L M A S R Q Q L A K E T S N F G F S L L R K I S M R H D G N M V F S P F G M S L A M T G L M L G A T G P T E T Q I
K R G L H L Q A L K P T K P G L L P S L F K G L R E T L S R N L E L G L S Q G S F A F I H K D F D V K E T F F N L S K R Y F
D T E C V P M N F R N A S Q A K R L M N H Y I N K E T R G K I P K L F D E I N P E T K L I L V D Y I L F K G K W L T P F D P
V F T E V D T F H L D K Y K T I K V P M M Y G A G K F A S T F D K N F R C H V L K L P Y Q G N A T M L V V L M E K M G D H L
A L E D Y L T T D L V E T W L R N M K T R N M E V F F P K F K L D Q K Y E M H E L L R Q M G I R R I F S P F A D L S E L S A
T G R N L Q V S R V L R R T V I E V D E R G T E A V A G I L S E I T A Y S M P P V I K V D R P F H F M I Y E E T S G M L L F
L G R V V N P T L L

296/330

FIGURE 296

MNQLSFLLFLIATTRGWSTDEANTYFKEWTCSSSPSLPRSCKEIKDECPSAFDGLYFLRTEN
GVIIYQTFCDMTSGGGGWTLVASVHENDMRGKCTVGDRWSSQQGSKADYPEGDGNWANYNTFG
SAEAATSDDYKNPGYYDIQAKDLGIWHVPNKSPMQHWRNSSLLRYRTDTGFLQTLGHNLFGI
YQKYPVKYGEKGCWTDNGPVI PVVYDFGDAQKTASYYS PYGQREFTAGFVQFRVFNNERAAN
ALCAGMRVTGCNTEHHCIGGGGYFPEAS PQQCGDFSGFDWSGYGTHVGYSSSREITEAAVLL
FYR

0993727 1100
10011 1216550

FIGURE 297

GCGGAGCCGGCGCCGGCTGCGCAGAGGAGCCGCTCTCGCCGCCGCCACCTCGGCTGGGAGCC
CACGAGGCTGCCGCATCCTGCCCTCGGAACA**ATG**GGACTCGGCGCGGAGGTGCTTGGGCCG
CGCTGCTCCTGGGGACGCTGCAGGTGCTAGCGCTGCTGGGGGCCGCCCATGAAAGCGCAGCC
ATGGCGGCATCTGCAAACATAGAGAATTCTGGGCTTCCACACAACCTCCAGTGCTAACTCAAC
AGAGACTCTCCAACATGTGCCTTCTGACCATACAAATGAACTTCCAACAGTACTGTGAAAC
CACCAACTTCAGTTGCCTCAGACTCCAGTAATACAACGGTCACCACCATGAAACCTACAGCG
GCATCTAATACAACAACACCAGGGATGGTCTCAACAAATATGACTTCTACCACCTTAAAGTC
TACACCCAAAACAACAAGTGTTTACAGAACACATCTCAGATATCAACATCCACAATGACCG
TAACCCACAATAGTTTCAGTGACATCTGCTGCTTCATCAGTAACAATCACAACAACCTATGCAT
TCTGAAGCAAAGAAAGGATCAAAATTTGATACTGGGAGCTTTGTTGGTGGTATTGTATTAAC
GCTGGGAGTTTTATCTATTCTTTACATTGGATGCAAAATGTATTACTCAAGAAGAGGCATTC
GGTATCGAACCATAGATGAACATGATGCCATCATT**TAA**GGAAATCCATGGACCAAGGATGGA
ATACAGATTGATGCTGCCCTATCAATTAATTTTGGTTTATTAATAGTTTAAACAATATTCT
CTTTTTGAAAATAGTATAAACAGGCCATGCATATAATGTACAGTGTATTACGTAAATATGTA
AAGATTCTTCAAGGTAACAAGGGTTTGGGTTTGAATAAACATCTGGATCTTATAGACCGT
TCATACAATGGTTTTAGCAAGTTCATAGTAAGACAAACAAGTCCTATCTTTTTTTTTTTGGCT
GGGGTGGGGGCATTGGTCACATATGACCAGTAATTGAAAGACGTCATCACTGAAAGACAGAA
TGCCATCTGGGCATACAAATAAGAAGTTTGTACAGCACTCAGGATTTTGGGTATCTTTTGT
AGCTCACATAAAGAACTTCAGTGCTTTTTCAGAGCTGGATATATCTTAATTACTAATGCCACA
CAGAAATTATACAATCAAACCTAGATCTGAAGCATAATTTAAGAAAAACATCAACATTTTTTG
TGCTTTAACTGTAGTAGTTGGTCTAGAAACAAAATACTCC

298/330

FIGURE 298

MGLGARGAWAALLLGTLQVLALLGAAHESAAMAASANIENSGLPHNSSANSTETLQHVP
SDH
TNETSNSTVKPPTSVASDSSNTTVTTMKPTAASNTTTPGMVSTNMTSTTLKSTPKTTSV
SQN
TSQISTSTMTVTHNSSVTSAASSVTITTTMHSEAKKGSKFDTGSFVGGIVLTLGVLSI
LYIG
CKMYYSRRGIRYRTIDEHDAII

bioRxiv preprint doi: <https://doi.org/10.1101/2020.04.14.038444>; this version posted April 14, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

300/330

FIGURE 300

MATLWGGLRLGSLLSLSCLALSVLLLAQLSDAAKNFEDVRCKCICPPYKENS
GHYKNIS
QKDCDCLHVVEPMPVRGPDVEAYCLRCECKYEERSSVTIKVTIIIIYLSILGLLLLYMVYLT
L
VEPILKRRLEFGHAQLIQSDDDIGDHQPFANAHDVLA
RSRANVLNKVEYAQQRWKLQVQEQ
RKSVFDRHVVL

090927 1400
105111 2268650

FIGURE 301

GCACCTGCGACCACCGTGAGCAGTCATGGCGTACTCCACAGTGCAGAGAGTCGCTCTGGCTT
CTGGGCTTGTCTGGCTCTGTCTGCTGCTGCCAAGGCCTTCCTGTCCCGCGGGAAGCGG
CAGGAGCCGCGCCGACACCTGAAGGAAAATTGGGCCGATTTCCACCTATGATGCATCATCA
CCAGGCACCCTCAGATGGCCAGACTCCTGGGGCTCGTTTCCAGAGGTCTCACCTTGCCGAGG
CATTTGCAAAGGCCAAAGGATCAGGTGGAGGTGCTGGAGGAGGAGGTAGTGGAAGAGGTCTG
ATGGGGCAGATTATTCCAATCTACGGTTTTGGGATTTTTTTATATATACTGTACATTCTATT
TAAGGTAAGTAGAATCATCCTAATCATATTACATCAATGA¹AAAATCTAATATGGCGATAAAAA
TCATTGTCTACATTAAACTTCTTATAGTTCATAAAATTATTTCAAATCCATCATCTCTTTA
AATCCTGCCTCCTCTTCATGAGGTACTTAGGATAGCCATTATTTAGTTTTACATAAGAATG
TTTACTCAATGTTTAAGTGTTTTGCCCCAAAATTCACAACTAACAAGGCAGAACTAGGACTT
GAACATGGATCTTTTGGTTCTTAATCCAGTGAGTGATAACAATTCAATGCACTCCCCTGCCA

302/330

FIGURE 302

MAYSTVQRVALASGLVLALSLLLPKAFLSRGKRQEPPTPEGKLGFRFPPMMHHHQAPSDGQT
PGARFQRSHLAEAFKAKGSGGGAGGGGSGRGLMGQIIPIYGFIFLYILYILFKVSRILI
ILHQ

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
1	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
2	2	4	6	8	10	12	14	16	18	20	22	24	26	28	30	32	34	36	38	40	42	44	46	48	50	52	54	56	58	60	62	64	66	68	70	72	74	76	78	80	82	84	86	88	90	92	94	96	98	100	102	104	106	108	110	112	114	116	118	120	122	124	126	128	130	132	134	136	138	140	142	144	146	148	150	152	154	156	158	160	162	164	166	168	170	172	174	176	178	180	182	184	186	188	190	192	194	196	198	200
3	3	6	9	12	15	18	21	24	27	30	33	36	39	42	45	48	51	54	57	60	63	66	69	72	75	78	81	84	87	90	93	96	99	102	105	108	111	114	117	120	123	126	129	132	135	138	141	144	147	150	153	156	159	162	165	168	171	174	177	180	183	186	189	192	195	198	201	204	207	210	213	216	219	222	225	228	231	234	237	240	243	246	249	252	255	258	261	264	267	270	273	276	279	282	285	288	291	294	297	300
4	4	8	12	16	20	24	28	32	36	40	44	48	52	56	60	64	68	72	76	80	84	88	92	96	100	104	108	112	116	120	124	128	132	136	140	144	148	152	156	160	164	168	172	176	180	184	188	192	196	200	204	208	212	216	220	224	228	232	236	240	244	248	252	256	260	264	268	272	276	280	284	288	292	296	300																									
5	5	10	15	20	25	30	35	40	45	50	55	60	65	70	75	80	85	90	95	100	105	110	115	120	125	130	135	140	145	150	155	160	165	170	175	180	185	190	195	200	205	210	215	220	225	230	235	240	245	250	255	260	265	270	275	280	285	290	295	300																																								
6	6	12	18	24	30	36	42	48	54	60	66																																																																																									

FIGURE 303

CGGCTCGAGTGCAGCTGTGGGGAGATTTTCAGTGCATTGCCTCCCCTGGGTGCTCTTCATCTT
GGATTTGAAAAGTTGAGAGCAGC**ATG**TTTTGCCCACTGAAACTCATCCTGCTGCCAGTGTTAC
TGGATTATTCCTTGGGCCTGAATGACTTGAATGTTTCCCCGCCTGAGCTAACAGTCCATGTG
GGTGATTACAGCTCTGATGGGATGTGTTTTCCAGAGCACAGAAGACAAATGTATATTCAAGAT
AGACTGGACTCTGTCACCAGGAGAGCACGCCAAGGACGAATATGTGCTATACTATTACTCCA
ATCTCAGTGTGCCTATTGGGCGCTTCCAGAACCGCGTACACTTGATGGGGGACATCTTATGC
AATGATGGCTCTCTCCTGCTCCAAGATGTGCAAGAGGCTGACCAGGGAACCTATATCTGTGA
AATCCGCCTCAAAGGGGAGAGCCAGGTGTTCAAGAAGGCGGTGGTACTGCATGTGCTTCCAG
AGGAGCCCAAAGAGCTCATGGTCCATGTGGGTGGATTGATTACAGATGGGATGTGTTTTCCAG
AGCACAGAAGTGAAACACGTGACCAAGGTAGAATGGATATTTTCAGGACGGCGCGCAAAGGA
GGAGATTGTATTTTCGTTACTACCACAAACTCAGGATGTCTGTGGAGTACTCCCAGAGCTGGG
GCCACTTCCAGAATCGTGTGAACCTGGTGGGGGACATTTTCCGCAATGACGGTTCCATCATG
CTTCAAGGAGTGAGGGAGTCAGATGGAGGAAACTACACCTGCAGTATCCACCTAGGGAACCT
GGTGTTCAAGAAAACCATTGTGCTGCATGTGAGCCCGGAAGAGCCTCGAACACTGGTGACCC
CGGCAGCCCTGAGGCCTCTGGTCTTGGGTGGTAATCAGTTGGTGATCATTGTGGGAATTGTC
TGTGCCACAATCCTGCTGCTCCCTGTTCTGATATTGATCGTGAAGAAGACCTGTGGAAATAA
GAGTTCAGTGAATTCTACAGTCTTGGTGAAGAACACGAAGAAGACTAATCCAGAGATAAAAG
AAAAACCCTGCCATTTTGAAAGATGTGAAGGGGAGAAACACATTTACTCCCCAATAATTGTA
CGGGAGGTGATCGAGGAAGAAGAACCAAGTGAAAAATCAGAGGCCACCTACATGACCATGCA
CCCAGTTTGGCCTTCTCTGAGGTCAGATCGGAACAACTCACTTGAAAAAAAGTCAGGTGGGG
GAATGCCAAAAACACAGCAAGCCTTTT**TGA**GAAAGAATGGAGAGTCCCTTCATCTCAGCAGCGG
TGGAGACTCTCTCCTGTGTGTGTCCTGGGCCACTCTACCAGTGATTTTCAGACTCCCGCTCTC
CCAGCTGTCTCCTGTCTCATTGTTTGGTCAATACACTGAAGATGGAGAATTTGGAGCCTGG
CAGAGAGACTGGACAGCTCTGGAGGAACAGGCCTGCTGAGGGGGAGGGGAGCATGGACTTGGC
CTCTGGAGTGGGACACTGGCCCTGGGAACCAGGCTGAGCTGAGTGGCCTCAAACCCCCCGTT
GGATCAGACCCTCCTGTGGGCAGGGTTCTTAGTGGATGAGTTACTGGGAAGAATCAGAGATA
AAAACCAACCCCAAATCAA

FIGURE 304

MFCPLKLILLPVLLDYSGLNDLNVSPPELTVHVGDSALMGCVFQSTEDKCIFKIDWTLSPG
EHAKDEYVLYYYSNLSVPIGRFQNRVHLMGDILCNDGSLLLQDVQEADQGTYICEIRLKGES
QVFKKAVVLHVLPEEPKELMVHVGGLIQMGCVFQSTEVKHVTKVEWIFSGRRAKEEIVFRYY
HKL RMSVEYSQSWGHFQNRVNLVGDI FRNDGSIMLQGVRES DGGNYTCSIHLGNLVFKKTIV
LHVSPEEPRTLVT PAALRPLVLGGNQLVII VGIVCATILLLPVLILIVKKTCGNKSSVNSTV
LVKNTKKTNPEIKEKPCHFERCEGEKHIYSPIIVREVIEEEEPSEKSEATYMTMHPVWPSLR
SDRNNSLEKKS GGGMPKTQQAF

304/330

FIGURE 305

CTATGAAGAAGCTTCCTGGAAAACAATAAGCAAAGGAAAACAAATGTGTCCCATCTCACATG
GTTCTACCCCTACTAAAGACAGGAAGATCATAAACTGACAGATACTGAAATTGTAAGAGTTGG
AAACTACATTTTGC AAAGTCATTGAACTCTGAGCTCAGTTGCAGTACTCGGGAAGCC**ATG**CA
GGATGAAGATGGATACATCACCTTAAATATTAAAACTCGGAAACCAGCTCTCGTCTCCGTTG
GCCCTGCATCCTCCTCCTGGTGGCGTGTGATGGCTTTGATTCTGCTGATCCTGTGCGTGGGG
ATGGTTGTCTGGGCTGGTGGCTCTGGGGATTTGGTCTGTCATGCAGCGCAATTACCTACAAGA
TGAGAATGAAAATCGCACAGGAAGCTCTGCAACAATTAGCAAAGCGCTTCTGTCAATATGTGG
TAAAACAATCAGAACTAAAGGGCACTTTC AAAGGTCATAAATGCAGCCCCTGTGACACAAAC
TGGAGATATTATGGAGATAGCTGCTATGGGTTCTTCAGGCACAACCTTAACATGGGAAGAGAG
TAAGCAGTACTGCACTGACATGAATGCTACTCTCCTGAAGATTGACAACCGGAACATTGTGG
AGTACATCAAAGCCAGGACTCATTTAATTCGTTGGGTCGGATTATCTCGCCAGAAGTCGAAT
GAGGTCTGGAAGTGGGAGGATGGCTCGGTTATCTCAGAAAATATGTTTGAGTTTTTGAAGA
TGGAAAAGGAAATATGAATTGTGCTTATTTTTCATAATGGGAAAATGCACCCTACCTTCTGTG
AGAACAAACATTATTTAATGTGTGAGAGGAAGGCTGGCATGACCAAGGTGGACCAACTACCT
TAATGCAAAGAGGTGGACAGGATAACACAGATAAGGGCTTTATTGTACAATAAAAGATATGT
ATGAATGCATCAGTAGCTGAAAAAAAAAAAAA

306/330

FIGURE 306

MQDEDGYITLNIKTRKPALVSVGPASSSWVRMALILLILCVGMVVGLVALGIWSVMQRNYL
QDENENRTGTLQQLAKRFCQYVVKQSELKGTFKGHKCSPCDTNWRYYGDSYGFRRHNLWE
ESKQYCTDMNATLLKIDNRNIVEYIKARTHLIRWVGLSRQKSNEVWKWEDGSVISENMFEFL
EDGKGNMNCAYFHNGKMHPTFCENKHLYLMCERKAGMTKVDQLP

306/330

CCCACGCGTCCGCGCAGTCTCGCGCAGTTCTGCCTCCGCTGCCAGTCTCTCGCCGCGATCCCCGG
CCCCGGGGCTGTGGCGTCTGACTCCGACCCAGGCAGCCAGCAGCCCCGCGCGGGAGCCGGACCCG
CGCCGGAGGAGCTCGGACGGCATGCTGAGCCCCCTCCTTTGCTGAAGCCCGAGTGCGGAGAA
GCCCCGGGCAAACGCAGGCTAAGGAGACCAAAGCGGCGAAGTTCGCGAGACAGCGGACAAGCAC
CGGAGGAGAAGGAGGAGGAGGCGAACCAGAGAGGGGCAGCAAAAGAAGCGGTGGTGGTGGG
CGTCGTGGCC**ATG**GCGGCGGCTATCGCCAGCTCGCTCATCCGTCAGAAGAGGCAAGCCCCGCG
AGCGCGAGAAATCCAACGCCTGCAAGTGTGTCAGCAGCCCCAGCAAAGGCAAGACCAGCTGC
GACAAAAACAAGTTAAATGTCTTTTCCGGGTCAAACCTCTTCGGCTCCAAGAAGAGGCGCAG
AAGAAGACCAGAGCCTCAGCTTAAGGGTATAGTTACCAAGCTATACAGCCGACAAGGCTACC
ACTTGCAGCTGCAGGCGGATGGAACCATTGATGGCACCAAAGATGAGGACAGCACTTACACT
CTGTTTAACCTCATCCCTGTGGGTCTGCGAGTGGTGGCTATCCAAGGAGTTCAAACCAAGCT
GTACTTGGCAATGAACAGTGAGGGATACTTGTACACCTCGGAACTTTTACACCTGAGTGCA
AATTCAAAGAATCAGTGTTTGAAAATTATTATGTGACATATTCATCAATGATATACCGTCAG
CAGCAGTCAGGCCGAGGGTGGTATCTGGGTCTGAACAAAGAAGGAGAGATCATGAAAGGCAA
CCATGTGAAGAAGAACAAGCCTGCAGCTCATTTTCTGCCTAAACCACTGAAAGTGGCCATGT
ACAAGGAGCCATCACTGCACGATCTCACGGAGTTCTCCCGATCTGGAAGCGGGACCCCAACC
AAGAGCAGAAGTGTCTCTGGCGTGCTGAACGGAGGCAAATCCATGAGCCACAATGAATCAAC
G**TAG**CCAGTGAGGGCAAAGAAGGGCTCTGTAACAGAACCTTACCTCCAGGTGCTGTTGAAT
TCTTCTAGCAGTCCTTCACCCAAAAGTTCAAATTTGTCAGTGACATTTACCAAACAAACAGG
CAGAGTTCACTATTCTATCTGCCATTAGACCTTCTTATCATCCATACTAAAGC

FIGURE 308

></usr/seqdb2/sst/DNA/Dnaseqs.full/ss.DNA28498

><subunit 1 of 1, 245 aa, 1 stop

><MW: 27564, pI: 10.18, NX(S/T): 1

MAAAIASSLIRQKRQAREREKSNACKCVSSPSKGKTSCDKNKLNVFSRVKLFSGSKRRRRRP
EPQLKGIVTKLYSRQGYHLQLQADGTIDGTKDEDSTYTLFNLIPVGLRVVAIQGVQTKLYLA
MNSEGILYTSELFTPECKFKESVFENYYVTYSSMIYRQQQSGRGWYLGLNKEGEIMKGNHVK
KNKPAAHFLPKPLKVAMYKEPSLHDLTEFSRSGSGTPTKRSVSGVLNGGKSMSHNEST

N-glycosylation site.

amino acids 242-246

Glycosaminoglycan attachment site.

amino acids 165-169, 218-222

Tyrosine kinase phosphorylation site.

amino acids 93-100

N-myristoylation site.

amino acids 87-93, 231-237

ATP/GTP-binding site motif A (P-loop).

amino acids 231-239

HBGF/FGF family proteins

amino acids 78-94, 102-153

310/330

FIGURE 310

MAALASSLIHQKREVREPGGSRPVSAQRRVCPRGTKSLCQKQLLILLSKVRLCGGRPARPDR
GPEPQLKGIVTKLFCRQGFYLLQANPDGSIQGTPEDTSSFTHFNLIPVGLRVVTIQSAKLGHY
MAMNAEGLLYSSPHFTAECRFKECVFENYYVLYASALYRQRRSGRAWYLGLDKEGQVMKGNR
VKKTKAAAHFLPKLLEVAMYQEPSLHSVPEASPSSPPAP

Tyrosine kinase phosphorylation site:

amino acids 199-207

N-myristoylation sites:

amino acids 54-60, 89-95, 131-137

HBGF/FGF family signature:

amino acids 131-155

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 314

MLNKMTLHPQQIMIGPRFNRALFDPLLVLALLQLLVVAGLVRAQTCPSVCSCSNQFSKVIC
 VRKNLREVPDGI STNTRLLNLHENQIQIIKVNSFKHLRHLEILQLSRNHIRTIEIGAFNGLA
 NLNTLELFDNRLTTIPNGAFVYLSKLELWLRNPNIESIPSYAFNRIPSLRRLDLGELKRLS
 YISEGAFEGLSNLRYNLAMCNLREIPNLTPLIKLDLDELDSGNHLSAIRPGSFQGLMHLQKL
 WMIQSQIQVIERNAFDNLQSLVEINLAHNNLTLLPHDLFTPLHHLERIHLLHNPWNCNDIL
 WLSWWIKDMAPSNTACCARCNTPPNLKGRYIGELDQNYFTCYAPVIVEPPADLNVTEGMAAE
 LKCRASSTLSVSWITPNGTVMTHGAYKVRIAVLSDGTNLNFTNVTVQDTGMYTCMVSNVSGN
 TTASATLNVTAATTTTFSYFSTVTVETMEPSQDEARTTDNNVGPTPVVDWETTNVTTSLTPQ
 STRSTEKTF TIPVTDINSGIPGIDEVMKTTKIIIGCFVAITLMAAVMLVIFYKMRKQHRQN
 HHAPTRTVEIINVDDEITGDTPMESHLPMPAIEHEHLNHYSYKSPFNHTTTVNTINSIHSS
 VHEPLLIRMNSKDNVQETQI

Signal sequence:

amino acids 1-44

Transmembrane domain:

amino acids 523-543

N-glycosylation site.

amino acids 278-282, 364-368, 390-394, 412-416, 415-419,
 434-438, 442-446, 488-492, 606-610

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 183-187

Casein kinase II phosphorylation site.

amino acids 268-272, 417-421, 465-469, 579-583, 620-624

N-myristoylation site.

amino acids 40-46, 73-79, 118-124, 191-197, 228-234, 237-243,
 391-397, 422-428, 433-439, 531-537

FIGURE 315

GCGCCGGAGCCCATCTGCCCCACAGGGGCACGGGGCGCGGGGCCGGCTCCCGCCCGGCACAT
GGCTGCAGCCACCTCGCGCGCACCCCGAGGCGCCGCGCCAGCTCGCCCGAGGTCCGTGCGA
GGCGCCCGGCCGCCCGGAGCCAAGCAGCAACTGAGCGGGGAAGCGCCCGCGTCCGGGGATC
GGG**ATG**TCCCTCCTCCTTCTCCTCCTTGCTAGTTTCCTACTATGTTGGAACCTTGGGGACTCA
CACTGAGATCAAGAGAGTGGCAGAGGAAAAGGTCACTTTGCCTGCCACCATCAACTGGGGC
TTCCAGAAAAAGACACTCTGGATATTGAATGGCTGCTCACCGATAATGAAGGGAACCAAAAA
GTGGTGATCACTTACTCCAGTCGTCATGTCTACAATAACTTGACTGAGGAACAGAAGGGCC
AGTGGCCTTTGCTTCCAATTTCTGCGCAGGAGATGCCTCCTTGAGATTGAACCTCTGAAGC
CCAGTGATGAGGGCCGGTACACCTGTAAGGTTAAGAATTCAGGGCGCTACGTGTGGAGCCAT
GTCATCTTAAAGTCTTAGTGAGACCATCCAAGCCCAAGTGTGAGTTGGAAGGAGAGCTGAC
AGAAGGAAGTGACCTGACTTTGCAGTGTGAGTCATCCTCTGGCACAGAGCCATTGTGTATT
ACTGGCAGCGAATCCGAGAGAAAGAGGGAGAGATGAACGTCTGCCTCCCAAATCTAGGATT
GACTACAACCACCTTGGACGAGTTCTGCTGCAGAATCTTACCATGTCCTACTCTGGACTGTA
CCAGTGCCACAGCAGGCAACGAAGCTGGGAAGGAAAGCTGTGTGGTGCGAGTAACGTACAGT
ATGTACAAAGCATCGGCATGGTTGCAGGAGCAGTGACAGGCATAGTGGCTGGAGCCCTGCTG
ATTTTCCTCTTGGTGTGGCTGCTAATCCGAAGGAAAGACAAAGAAAGATATGAGGAAGAAGA
GAGACCTAATGAAATTCGAGAAGATGCTGAAGCTCCAAAAGCCCGTCTTGTGAAACCCAGCT
CCTCTTCTCAGGCTCTCGGAGCTCACGCTCTGGTTCTTCTCCTCCACTCGCTCCACAGCAAAT
AGTGCTCACGCAGCCAGCGGACACTGTCAACTGACGCAGCACCCAGCCAGGGCTGGCCAC
CCAGGCATACAGCCTAGTGGGGCCAGAGGTGAGAGGTTCTGAACCAAAGAAAGTCCACCATG
CTAATCTGACCAAAGCAGAAACCACACCCAGCATGATCCCCAGCCAGAGCAGAGCCTTCCAA
ACGGTCT**TGA**ATTACAATGGACTTGACTCCACGCTTTCTTAGGAGTCAGGGTCTTTGGACTC
TTCTCGTCATTGGAGCTCAAGTCACCAGCCACACAACCAGATGAGAGGTCATCTAAGTAGCA
TTGAGCATTCACGGAACAGATTGAGATGAGCATTTTCTTATACAATACCAAACAAGCAAA
AGGATGTAAGCTGATTCATCTGTAAAAAGGCATCTTATTGTGCCTTTAGACCAGAGTAAGG
AAAGCAGGAGTCCAAATCTATTTGTTGACCAGGACCTGTGGTGAGAAGGTTGGGGAAAGGTG
AGGTGAATATACCTAAAACCTTTTAATGTGGGATATTTGTATCAGTGCTTTGATTCACAATT
TTCAAGAGGAAATGGGATGCTGTTTGTAAATTTTCTATGCATTTCTGCAAACCTTATTGGATT
ATTAGTTATTGAGACAGTCAAGCAGAACCCACAGCCTTATTACACCTGTCTACACCATGTAC
TGAGCTAACCACTTCTAAGAACTCCAAAAAAGGAAACATGTGTCTTCTATTCTGACTTAAC
TTCATTTGTCATAAGGTTTGGATATTAATTTCAAGGGGAGTTGAAATAGTGGGAGATGGAGA
AGAGTGAATGAGTTTCTCCCACTCTATACTAATCTCACTATTTGTATTGAGCCCCAAAATAAC
TATGAAAGGAGACAAAAATTTGTGACAAAGGATTGTGAAGAGCTTTCCATCTTCATGATGTT
ATGAGGATTGTTGACAAACATTAGAAATATATAATGGAGCAATTGTGGATTTCCCCCTCAAAT
CAGATGCCTCTAAGGACTTTCTGCTAGATATTTCTGGAAGGAGAAAATACAACATGTCATT
TATCAACGTCCTTAGAAAGAATTCTTCTAGAGAAAAAGGGATCTAGGAATGCTGAAAGATTA
CCCAACATACCATTTATAGTCTCTTCTTTCTGAGAAAAATGTGAAACCAGAATTGCAAGACTGG
GTGGACTAGAAAGGGAGATTAGATCAGTTTTCTCTTAATATGTCAAGGAAGGTAGCCGGGCA
TGGTGCCAGGCACCTGTAGGAAAATCCAGCAGGTGGAGGTTGCAGTGAGCCGAGATTATGCC
ATTGCACTCCAGCCTGGGTGACAGAGCGGGACTCCGTCTC

316/330

FIGURE 316

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA45419

><subunit 1 of 1, 373 aa, 1 stop

><MW: 41281, pI: 8.33, NX(S/T): 3

MSLLLLLLLLVSYYVGTLGTHTEIKRVAEEKVTLPCHHQLGLPEKDTLDIEWLLTDNEGNQKV
VITYSSRHVYNNLTTEEQKGRVAFASNFLAGDASLQIEPLKPSDEGRYTCKVKNSGRYVWSHV
ILKVLVRPSKPKCELEGELTEGSDLTLCESSSGTEPIVYYWQRIREKEGEDERLPPKSRID
YNHPGRVLLQNLTMSSGLYQCTAGNEAGKESCVVRVTVQYVQSIGMVAGAVTGIVAGALLI
FLLVWLLIRRKDKERYEEEEERPNEIREDAEAPKARLVKPSSSSSGSRSSRSGSSSTRSTANS
ASRSQRTLSTDAAPQPGLATQAYSLVGPEVRGSEPKKVHHANLTKAETTPSMIPSQSRAFQTV

Signal sequence:

amino acids 1-16

Transmembrane domain:

amino acids 232-251

FIGURE 317

CGCGAGGCGCGGGGAGCCTGGGACCAGGAGCGAGAGCCGCTACCTGCAGCCGCCGCCACGGCACGGCAGCCA
 CCATGCGGCTCCTGCTGTGCTTCGTGCTCCTGTGCGGAGTAGTGGATTTGCGCCAGAAGTTTGGAGTATCACTACT
 CCTGAAGAGATGATTGAAAAAGCCAAAGGGGAACTGCCTATCTGCCATGCAAATTTACGCTTAGTCCCGAAGA
 CCAGGGACCGCTGGACATCGAGTGGCTGATATCACCAGCTGATAATCAGAAGGTGGATCAAGTGATTATTTTAT
 ATTCTGGAGACAAAATTTATGATGACTACTATCCAGATCTGAAAGGCCGAGTACATTTTACGAGTAATGATCTC
 AAATCTGGTGATGCATCAATAAATGTAACGAATTTACAACGTGTGAGATATTGGCACATATCAGTGCAAAGTGAA
 AAAAGCTCCTGGTGTGCAAATAAGAAGATTCATCTGGTAGTTCCTTGTAAAGCCTTCAGGTGCGAGATGTTACG
 TTGATGGATCTGAAGAAATTGGAAGTGACTTTAAGATAAAATGTGAACCAAAAGAAGGTTCACTTCCATTACAG
 TATGAGTGGCAAAAATTTGCTGACTCACAGAAAATGCCCACTTCATGGTTAGCAGAAATGACTTCATCTGTTAT
 ATCTGTAAAAAATGCCCTCTTCTGAGTACTCTGGGACATACAGCTGTACAGTCAGAAACAGAGTGGGCTCTGATC
 AGTGCCCTGTTGCGTCTAAACGTTGTCCCTCCTTCAAATAAAGCTGGACTAATTGCAGGAGCCATTATAGGAAC
 TTGCTTGCTCTAGCGCTCATTGGTCTTATCATCTTTTGTGTCGTAAAAAGCGCAGAGAAGAAAAATATGAAAA
 GGAAGTTCATCACGATATCAGGGAAGATGTGCCACCTCCAAAGAGCCGTACGTCCACTGCCAGAAGCTACATCG
 GCAGTAATCATTCATCCCTGGGGTCCATGTCTCCTTCCAACATGGAAGGATATTCCAAGACTCAGTATAACCAA
 GTACAGTGAAGACTTTGAACGCACTCCTCAGAGTCCGACTCTCCACCTGCTAAGTTCAGTACCCTTACAA
 GACTGATGGAATTACAGTTGTATATAATATGGACTACTGAAGAATCTGAAGTATTGTATTATTTGACTTTATTTT
 AGGCCTCTAGTAAAGACTTAAATGTTTTTAAAAAAGCACAAAGGCACAGAGATTAGAGCAGCTGTAAGAACAC
 ATCTACTTTTATGCAATGGCATTAGACATGTAAGTCAGATGTCATGTCAAATTTAGTACGAGCCAAATCTTTGT
 TAAAAACCCTATGTATAGTGACACTGATAGTTAAAAGATGTTTTATTATATTTTCAATAACTACCACTAACAA
 ATTTTAACTTTTCATATGCATATTTCTGATATGTGGTCTTTTAGGAAAAGTATGGTTAATAGTTGATTTTCAA
 AGGAAATTTTAAATTTCTACGTTCTGTTTAAATGTTTTGCTATTTAGTTAAATACATTGAAGGGAAATACCCG
 TTCTTTTCCCCTTTTATGCACACAACAGAAACACGCGTGTGTCATGCCCTCAAACATTTTTTTATTTGCAACTACA
 TGATTTTACACAATTTCTTAAACAACGACATAAAATAGATTTCTTGTATATAAATAACTTACATACGCTCCA
 TAAAGTAAATTTCTCAAAGGTGCTAGAACAAATCGTCCACTTCTACAGTGTTCTCGTATCCAACAGAGTTGATGC
 ACAATATATAAATACTCAAGTCCAATATTAATAAATTTAGGCACTTGACTAACTTTAATAAATTTCTCAAAC
 TATCAATATCTAAAGTGATATATTTTTTAAAGAAAGATTATTTCTCAATAACTTCTATAAAAAATAAGTTTGATGG
 TTTGGCCCATCTAACTTCACTACTATTAGTAAGAATTTTAACTTTTAAATGTGTAGTAAGGTTTATTCTACCTT
 TTTCTCAACATGACACCAACACAATCAAAAACGAAGTTAGTGAGGTGCTAACATGTGAGGATTAATCCAGTGAT
 TCCGGTCACAATGCATTTCCAGGAGGAGGTACCCATGTCACTGGAATTGGGCGATATGGTTTATTTTTCTTCCC
 TGATTTGGATAACCAAATGGAACAGGAGGAGGATAGTGATTCTGATGGCCATTCCCTCGATACATTCCTGGCTT
 TTTTCTGGGCAAAGGTGCCACATTTGGAAGAGGTGGAATATAAGTTCTGAAATCTGTAGGGAAGAGAACACAT
 TAAGTTAATTTCAAAGGAAAAAATCATCATCTATGTTCCAGATTTCTCATTAAAGACAAAGTTACCCACAACACT
 GAGATCACATCTAAGTGACACTCCTATTGTGAGGTCTAAATACATTAAAAACCTCATGTGTAATAGGCGTATAA
 TGTATAACAGGTGACCAATGTTTTCTGAATGCATAAAGAAATGAATAAACTCAAACACAGTACTTCTTAAACAA
 CTTCAACCAAAAAAGACCAAAACATGGAACGAATGGAAGCTTGTAAGGACATGCTTGTTTTAGTCCAGTGTTTT
 CCACAGCTGGCTAAGCCAGGAGTCACTTGGAGGCTTTTAAATACAAAACATTGGAGCTGGAGGCCATTATCCTT
 AGCAAACATAATGCAGAAACAGAAAATCAACTACCGCATGTTCTCACTTATAAGTGGGAGGTAATGATAAGA
 TATGAACACAAAGAAGGAAACAATAGACATTGGAGTCTATTTGAGAGGGGAGGTGGGAGAAGGAAAAGGAGCA
 GAAAAGATAACTATTGAGTACTGCCCTTACACCTGGGTGATGAAATAATATGTACAACAAATCCCTGTGACACA
 TGTTTACCTATGGAACAAACCTTCATGTGTATCCCTAAACCTAAAAATAAAAGTTAAAAA
 AA
 AA

FIGURE 318

```
></usr/seqdb2/sst/DNA/Dnaseqs.full/ss.DNA82361
```

```
><subunit 1 of 1, 352 aa, 1 stop
```

><MW: 38938, pI: 7.86, NX(S/T): 3

MALLLCFVLLCGVVDFARSLSITTPPEEMIEKAKGETAYLPCKFTLSPEDQGPLDIEWLISPA
DNQKVDQVIILYSGDKIYDDYYPDLKGRVHFTSNDLKSGDASINVTNLQLSDIGTYQCKVKK
APGVANKKIHVLVLVKPSGARCIVDGGSEEIGSDFKIKCEPKEGSLPLQYEWQKLSDSQKMPT
SWLAEMTSSVISVKNASSEYSGTYSCTVRNRVGSQCLLRLNVVPPSNKAGLIAGAIIGTLL
ALALIGLIIFFCCRKKRREEKYEKEVHHDIREDVPPPKSRTSTARSYIGSNHSSLGSMSPSNM
EGYSKTOYNOVPSSEDFERTPOSPTLPPAKFKYPYKTDGITVV

Signal sequence.

amino acids 1-19

Transmembrane domain:

amino acids 236-257

N-glycosylation sites.

amino acids 106-110, 201-205, 298-302

Tyrosine kinase phosphorylation sites.

amino acids 31-39, 78-85, 262-270

N-myristoylation sites.

amino acids 116-122, 208-214, 219-225, 237-243, 241-247,
245-251, 296-302

Myelin P0 protein.

amino acids 96-125

FIGURE 319

TGAATGACTTCCACGGCTGGGACGGGAACCTTCCACCCACAGCTATGCCTCTGATTGGTGA
 ATGGTGAAGGTGCCTGTCTAACTTTTCTGTAAAAAGAACCAGCTGCCTCCAGGCAGCCAGCC
 CTCAAGCATCACTTACAGGACCAGAGGGACAAGACATGACTGTGATGAGGAGCTGCTTTCGC
 CAATTTAACACCAAGAAGAATTGAGGCTGCTTGGGAGGAAGGCCAGGAGGAACACGAGACTG
 AGAG**ATGA**ATTTTCAACAGAGGCTGCAAAGCCTGTGGACTTTAGCCAGACCCTTCTGCCCTC
 CTTTGGCTGGCGACAGCCTCTCAAATGCAGATGGTTGTGCTCCCTTGCCTGGGTTTTACCCTG
 CTTCTCTGGAGCCAGGTATCAGGGGCCAGGGCCAAGAATCCACTTTGGGGCCCTGCCAAGT
 GAAGGGGGTTGTTCCCCAGAACTGTGGGAAGCCTTCTGGGCTGTGAAAGACACTATGCAAG
 CTCAGGATAACATCACGAGTGCCCGGCTGCTGCAGCAGGAGGTTCTGCAGAACGTCTCGGAT
 GCTGAGAGCTGTTACCTTGTCCACACCCTGCTGGAGTTCTACTTGAAAACGTTTTTCAAAAA
 CCACCACAATAGAACAGTTGAAGTCAGGACTCTGAAGTCATTCTCTACTCTGGCCAACAAC
 TTGTTCTCATCGTGTCAAACTGCAACCCAGTCAAGAAAATGAGATGTTTTCCATCAGAGAC
 AGTGCACACAGGCGGTTTCTGCTATTCCGGAGAGCATTCAAACAGTTGGACGTAGAAGCAGC
 TCTGACCAAAGCCCTTGGGGAAGTGGACATTCTTCTGACCTGGATGCAGAAATTCTACAAGC
 TC**TGA**ATGTCTAGACCAGGACCTCCCTCCCCCTGGCACTGGTTTGTTCCTGTGTCATTTCA
 AACAGTCTCCCTTCTTATGCTGTTCACTGGACACTTCACGCCCTTGGCCATGGGTCCCATT
 TTGGCCCAGGATTATTGTCAAAGAAGTCATTCTTTAAGCAGCGCCAGTGACAGTCAGGGAAG
 GTGCCTCTGGATGCTGTGAAGAGTCTACAGAGAAGATTCTTGTATTTATTACAACCTCTATTT
 AATTAATGTGAGTATTTCAACTGAAGTTCTATTTATTTGTGAGACTGTAAGTTACATGAAGG
 CAGCAGAATATTGTGCCCCATGCTTCTTTACCCCTCACAATCCTTGCCACAGTGTGGGGCAG
 TGGATGGGTGCTTAGTAAGTACTTAATAAACTGTGGTGCTTTTTTTTGGCCTGTCTTTGGATT
 GTTAAAAAACAGAGAGGGATGCTTGGATGTAAAACTGAACTTCAGAGCATGAAAATCACACT
 GTCTTCTGATATCTGCAGGGACAGAGCATTGGGGTGGGGGTAAGGTGCATCTGTTTGAAAAG
 TAAACGATAAAAATGTGGATTAAAGTGCCCGAGCACAAAGCAGATCCTCAATAAACATTTTATT
 TCCCACCCACACTCGCCAGCTCACCCCATCATCCCTTTCCCTTGGTGCCCTCCTTTTTTTTTT
 TATCCTAGTCATTCTTCCCTAATCTTCCACTTGAGTGTCAAGCTGACCTTGCTGATGGTGAC
 ATTGCACCTGGATGTACTATCCAATCTGTGATGACATTCCCTGCTAATAAAAGACAACATAA
 CTCCAAAAA

320/330

FIGURE 320

></usr/seqdb2/sst/DNA/Dnaseqs.full/ss.DNA88002

><subunit 1 of 1, 206 aa, 1 stop

><MW: 23799, pI: 9.12, NX(S/T): 3

MNFQQRLQSLWTLARPFPCPLLATASQMQMVVLPCLGFTLLLWSQVSGAQGQEFHFGPCQVK
GVVPQKLWEAFWAVKDTMQAQDNITSARLLQQEVLQNVSDAESCYLEVHTLLEFYLKTVFKNH
HNRTVEVRTLKSFSTLANNFVLIVSQLQPSQENEMFSIRDSAHRRLFFRAFKQLDVEAAL
TKALGEVDILLTWMQKFKYL

Signal sequence:

amino acids 1-42

N-glycosylation sites.

amino acids 85-89, 99-103, 126-130

protein

AAGGAGCAGCCCCGCAAGCACCAAGTGAGAGGCATGAAAGTTACAGTGTGTTTCCCTTTGGCTC
CTGGGTACAATACTGATATTGTGCTCAGTAGACAACCACGGTCTCAGGAGATGTCTGATTTCC
CACAGACATGCACCATATAGAAGAGAGTTTCCAAGAAATCAAAAGAGCCATCCAAGCTAAGG
ACACCTTCCCAAATGTCACCTATCCTGTCCACATTGGAGACTCTGCAGATCATTAAGCCCTTA
GATGTGTGCTGCGTGACCAAGAACCTCCTGGCGTTCTACGTGGACAGGGTGTTCAAGGATCA
TCAGGAGCCAAACCCCAAAATCTTGAGAAAAATCAGCAGCATTGCCAACTCTTTCCTCTACA
TGCAGAAACTCTGCGGCAATGTCAGGAACAGAGGCAGTGTCACTGCAGGCAGGAAGCCACC
AATGCCACCAGAGTCATCCATGACAACCTATGATCAGCTGGAGGTCCACGCTGCTGCCATTAA
ATCCCTGGGAGAGCTCGACGTCTTTCTAGCCTGGATTAATAAGAATCATGAAGTAATGTTCT
CAGCTTGATGACAAGGAACCTGTATAGTGATCCAGGGATGAACACCCCCTGTGCGGTTTACT
GTGGGAGACAGCCCACCTTGAAGGGGAAGGAGATGGGGAAGGCCCTTGCAGCTGAAAGTCC
CACTGGCTGGCCTCAGGCTGTCTTATTCCGCTTGAAAATAGGC AAAAAGTCTACTGTGGTAT
TTGTAATAAACTCTATCTGCTGAAAGGGCCTGCAGGCCATCCTGGGAGTAAAGGGCTGCCTT
CCCATCTAATTTATTGTAAAGTCATATAGTCCATGTCTGTGATGTGAGCCAAGTGATATCCT
GTAGTACACATTGTACTGAGTGGTTTTTCTGAATAAATTCCATATTTTACCTATGA

FIGURE 322

></usr/seqdb2/sst/DNA/Dnaseqs.full/ss.DNA92282
><subunit 1 of 1, 177 aa, 1 stop
><MW: 20452, pI: 8.00, NX(S/T): 2
MKLQCVSLWLLGTILILCSVDNHGLRRCLISTDMHHIEESFQEIKRAIQAKDTFPNVTILST
LETLQIIKPLDVCCVTKNLLAFYVDRVFKDHQEPNPKILRKISSIANSFLYMQKTLRQCQEQ
RQCHCRQEATNATRVIHNDNYDQLEVHAAAIKSLGELDVFLAWINKNHEVMFSA

Signal sequence:

amino acids 1-18

N-glycosylation sites.

amino acids 56-60, 135-139

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 102-106

N-myristoylation site.

amino acids 24-30

Actinin-type actin-binding domain signature 1.

amino acids 159-169

FIGURE 323

CCCGTGCCCAAGAGTGACGTAAAGTACCGCCTATAGAGTCTATAGGCCCACTTGGGCTTCGTTAG
AACGCGGCTACAATTAATACATAACCTTATGTATCATACACATACGATTTAGGTGACACTAT
AGAATAACATCCACTTTGCCTTTCTCTCCACAGGTGTCCACTCCCAGGTCCAACCTGCACCTC
GGTTCTATCGATAATCTCAGCACCAGCCACTCAGAGCAGGGCACGATGTTGGGGGCCCCGCCT
CAGGCTCTGGGTCTGTGCCTTGTGCAGCGTCTGCAGCATGAGCGTCCTCAGAGCCTATCCCCA
ATGCCTCCCCACTGCTCGGCTCCAGCTGGGGTGGCCTGATCCACCTGTACACAGCCACAGCC
AGGAACAGCTACCACCTGCAGATCCACAAGAATGGCCATGTGGATGGCGCACCCCATCAGAC
CATCTACAGTGCCCTGATGATCAGATCAGAGGATGCTGGCTTTGTGGTGATTACAGGTGTGA
TGAGCAGAAGATACCTCTGCATGGATTTTCAGAGGCAACATTTTTTGGATCACACTATTTTCGAC
CCGGAGAACTGCAGGTTCCAACACCAGACGCTGGAAAACGGGTACGACGTCTACCACTCTCC
TCAGTATCACTTCCTGGTCAGTCTGGGGCCGGGCGAAGAGAGCCTTCCTGCCAGGCATGAACC
CACCCCCGTA CTCCCAGTTCCTGTCCCGGAGGAACGAGATCCCCCTAATTCACTTCAACACC
CCCATACCACGGCGGCACACCCGGAGCGCCGAGGACGACTCGGAGCGGGACCCCCTGAACGT
GCTGAAGCCCCGGGCCCCGGATGACCCCGGCCCCGGCCTCCTGTTACAGGAGCTCCCGAGCG
CCGAGGACAACAGCCCGATGGCCAGTGACCCATTAGGGGTGGTCAGGGGCGGTTCGAGTGAAC
ACGCACGCTGGGGGAACGGGCCCCGAAGGCTGCCGCCCCCTTCGCCAAGTTCATCTAGGGTCTG
CTGG

FIGURE 324

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA142238

><subunit 1 of 1, 251 aa, 1 stop

><MW: 27954, pI: 9.22, NX(S/T): 1

MLGARLRLWVCALCSVCSMSVLRAYPNASPLLGSWSGGLIHLYTATARNSYHLQIHKNGHVD
GAPHQTIYSALMIRSEDAGFVVITGVMSRRYLCLMDFRGNI FGSHYFDPENCRFQHQTLENGY
DVIYHSPQYHFLVSLGRAKRAFLPGMNPPPYSQFLSRRNEIPLIHFNTPIPRRHTRSAEDDSE
RDPLNVLKPRARMTPAPASCSQELPSAEDNSPMASDPLGVVRGGRVNT HAGGTGPEGCRPFA
KFI

Important features of the protein:

Signal peptide:

amino acids 1-24

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 175-179

N-myristoylation site.

amino acids 33-39, 100-106, 225-231, 229-235

HBGF/FGF family proteins

amino acids 73-124

FIGURE 325

GGAAAAGGTACCCGCGAGAGACAGCCAGCAGTTCTGTGGAGCAGCGGTGGCCGGCTAGG**ATG**
 GGCTGTCTCTGGGGTCTGGCTCTGCCCCTTTTCTTCTTCTGCTGGGAGGTTGGGGTCTCTGG
 GAGCTCTGCAGGCCCCAGCACCCGCGAGAGCAGACACTGCGATGACAACGGACGACACAGAAG
 TGCCCGCTATGACTCTAGCACCGGGCCACGCCGCTCTGGAAACTCAAACGCTGAGCGCTGAG
 ACCTCTTCTAGGGCCTCAACCCCAGCCGGCCCCATTCCAGAAGCAGAGACCAGGGGAGCCAA
 GAGAATTTCCCCTGCAAGAGAGACCAGGAGTTTCACAAAACATCTCCCAACTTCATGGTG
 TGATCGCCACCTCCGTGGAGACATCAGCCGCCAGTGGCAGCCCCGAGGGAGCTGGAATGACC
 ACAGTTCAGACCATCACAGGCAGTGATCCCGAGGAAGCCATCTTTGACACCCTTTGCACCGA
 TGACAGCTCTGAAGAGGCCAAAGACACTCACAATGGACATATTGACATTGGCTCACACCTCCA
 CAGAAGCTAAGGGCCTGTCCTCAGAGAGCAGTGCCTCTTCCGACGGCCCCCATCCAGTCATC
 ACCCCGTACGGGCCTCAGAGAGCAGCGCCTCTTCCGACGGCCCCCATCCAGTCATCACCCC
 GTCACGGGCCTCAGAGAGCAGCGCCTCTTCCGACGGCCCCCATCCAGTCATCACCCCGTCAT
 GGTCCCCGGGATCTGATGTCACCTCTCCTCGCTGAAGCCCTGGTGACTGTCAAAACATCGAG
 GTTATTAATTGCAGCATCACAGAAATAGAAACAACAACTTCCAGCATCCCTGGGGCCTCAGA
 CATAGATCTCATCCCCACGGAAGGGGTGAAGGCCTCGTCCACCTCCGATCCACCAGCTCTGC
 CTGACTCCACTGAAGCAAACACACATCACTGAGGTCACAGCCTCTGCCGAGACCCTGTCC
 ACAGCCGGCACCACAGAGTCAGCTGCACCTCATGCCACGGTTGGGACCCCCACTCCCCACTAA
 CAGCGCCACAGAAAGAGAAGTGACAGCACCCGGGGCCACGACCCTCAGTGGAGCTCTGGTCA
 CAGTTAGCAGGAATCCCCTGGAAGAAACCTCAGCCCTCTCTGTTGAGACACCAAGTTACGTC
 AAAGTCTCAGGAGCAGCTCCGGTCTCCATAGAGGCTGGGTGAGCAGTGGGCAAAACAACCTTC
 CTTTGCTGGGAGCTCTGCTTCCTCCTACAGCCCCTCGGAAGCCGCCCTCAAGAACTTCACCC
 CTTCAGAGACACCGACCATGGACATCGCAACCAAGGGGCCCTTCCCCACCAGCAGGGACCCT
 CTTCCCTTCTGTCCCTCCGACTACAACCAACAGCAGCCGAGGGACGAACAGCACCTTAGCCAA
 GATCACAACCTCAGCGAAGACCACGATGAAGCCCCAACAGCCACGCCACGACTGCCCGGAC
 GAGGCCGACCACAGACG**TGA**GTGCAGGTGAAAATGGAGGTTTCCTCCTCCTGCGGCTGAGTG
 TGGCTTCCCCGGAAGACCTCACTGACCCAGAGTGGCAGAAAGGCTGATGCAGCAGCTCCAC
 CGGGAACCTCCACGCCACGCGCCTCACTTCCAGGTCTCCTTACTGCGTGTCAGGAGAGGCTA
 ACGGACATCAGCTGCAGCCAGGCATGTCCCGTATGCCAAAAGAGGGTGCTGCCCCTAGCCTG
 GGCCCCCACCAGACAGACTGCAGCTGCGTTACTGTGCTGAGAGGTACCCAGAAGGTTCCCATG
 AAGGGCAGCATGTCCAAGCCCCTAACCCAGATGTGGCAACAGGACCCTCGCTCACATCCAC
 CGGAGTGTATGTATGGGGAGGGGCTTCACCTGTTCCAGAGGTGTCCTTGGACTCACCTTGG
 CACATGTTCTGTGTTTCAGTAAAGAGAGACCTGATCACCCATCTGTGTGCTTCCATCCTGCA
 TTAAATTCACCTCAGTGTGGCCCCAAAAAAA

FIGURE 326

MGCLWGLALPLFFFCWEVGVSGSSAGPSTRADTAMTTDDTEVPAMTLAPGHAALETQTL
ETSSRASTPAGPIPEAETRGAKRISPAETRSFTKTSPNFMVLIATSVETSAASGSPEGAGM
TTVQTITGSDPEEAI FDTLCTDDSSSEEAKTLTMDILTLAHTSTEAKGLSSESSASSDGPHPV
ITPSRASESSASSDGPHPVITPSRASESSASSDGPHPVITPSWSPGSDVTLLEALVTVTNI
EVINCSITEIETTTSSIPGASDIDLIPTEGVKASSTSDPPALPDSTEAKPHITEVTASAETL
STAGTTESAAPHATVGTPLPTNSATEREVTAPGATTLGALVTVSRNPLEETSALSVETPSY
VKVSGAAPVSIEAGSAVGKTTSFAGSSASSYSPSEAALKNFTPSETPTMDIATKGPFPSTRD
PLPSVPPTTTNSSRGNTSLAKITTSAKTTMKPQQPRPRLPGRGRPQT

N-glycosylation sites:

amino acids 252-256, 445-449, 451-455

cAMP-and cGMP-dependent protein kinase phosphorylation site.

amino acids 84-90

Casein kinase II phosphorylation sites.

amino acids 37-41, 108-112, 131-135, 133-137, 148-152, 165-169,
246-250, 254-258, 256-260, 269-273, 283-287, 333-337, 335-339,
404-408, 414-418, 431-435

N-myristoylation sites.

amino acids 2-8, 19-25, 117-123, 121-127, 232-238, 278-284, 314-
320, 349-355, 386-392, 397-403, 449-455

ATP/GTP-binding site motif A (P-loop).

amino acids 385-393

FIGURE 327

GCGGAGCATCCGCTGCGGTCTCGCCGAGACCCCGCGCGGATTGCGCCGGTCTTCCCGCGG
 GCGCGACAGAGCTGTCCTCGCACCTGGATGGCAGCAGGGGGCGCCGGGGTCTCTCGACGCCA
 GAGAGAAATCTCATCATCTGTGCAGCCTTCTTAAAGCAAACCTAAGACCAGAGGGAGGATTAT
 CCTTGACCTTTGAAGACCAAACTAACTGAAATTTAAA**ATG**TTCCTTCGGGGGAGAAGGGAG
 CTTGACTTACACTTTGGTAATAATTTGCTTCCTGACACTAAGGCTGTCTGCTAGTCAGAATT
 GCCTCAAAAAGAGTCTAGAAGATGTTGTCATTGACATCCAGTCATCTCTTTCTAAGGGAATC
 AGAGGCAATGAGCCCGTATATACTTCAACTCAAGAAGACTGCATTAATTCCTTGCTGTTCAAC
 AAAAAACATATCAGGGGACAAAGCATGTAACCTTGATGATCTTCGACACTCGAAAAACAGCTA
 GACAACCCAACCTGCTACCTATTTTTCTGTCCCAACGAGGAAGCCTGTCCATTGAAACCAGCA
 AAAGGACTTATGAGTTACAGGATAATTACAGATTTTCCATCTTTGACCAGAAATTTGCCAAG
 CCAAGAGTTACCCCAGGAAGATTCTCTCTTACATGGCCAATTTTCACAAGCAGTCACTCCCC
 TAGCCCATCATCACACAGATTATTCAAAGCCACCGATATCTCATGGAGAGACACACTTTCT
 CAGAAGTTTGGATCCTCAGATCACCTGGAGAACTATTTAAGATGGATGAAGCAAGTGCCCA
 GCTCCTTGCTTATAAGGAAAAAGGCCATTCTCAGAGTTCACAATTTTCTCTGATCAAGAAA
 TAGCTCATCTGCTGCCTGAAAATGTGAGTGCCTCCAGCTACGGTGGCAGTTGCTTCTCCA
 CATAACACCTCGGCTACTCCAAAGCCCGCCACCCTTCTACCCACCAATGCTTCAGTGACACC
 TTCTGGGACTTCCCAGCCACAGCTGGCCACCACAGCTCCACCTGTAACCACTGTCACCTCTC
 AGCCTCCACGACCCTCATTTCTACAGTTTTTTACACGGGCTGCGGCTACACTCCAAGCAATG
 GCTACAACAGCAGTTCTGACTACCACCTTTCAGGCACCTACGGACTCGAAAGGCAGCTTAGA
 AACCATACCGTTTACAGAAATCTCCAACCTTAACCTTTGAACACAGGGAATGTGTATAACCCTA
 CTGCACTTTTCTATGTCAAATGTGGAGTCTTCCACTATGAATAAACTGCTTCCTGGGAAGGT
 AGGGAGGCCAGTCCAGGCAGTTCTTCCAGGGCAGTGTTCCAGAAAATCAGTACGGCCTTCC
 ATTTGAAAAATGGCTTCTTATCGGGTCCCTGCTCTTTGGTGTCTGTTCTGCTGGTGATAGGCC
 TCGTCCCTCCTGGGTAGAATCCTTTGGAATCACTCCGCAGGAAACGTTACTCAAGACTGGAT
 TATTTGATCAATGGGATCTATGTGGACATCT**TAAG**GATGGAACCTCGGTGTCTCTTAATTCATT
 TAGTAACCAGAAGCCCAAATGCAATGAGTTTCTGCTGACTTGCTAGTCTTAGCAGGAGGTG
 TATTTTGAAGACAGGAAAATGCCCCCTTCTGCTTTCCTTTTTTTTTTTTGGAGACAGAGTCTT
 GCTCTGTTGCCAGGCTGGAGTGCAGTAGCACGATCTCGGCTCTCACCGCAACCTCCGTCTC
 CTGGGTTCAAGCGATTCTCCTGCCTCAGCCTCCTAAGTATCTGGGATTACAGGCATGTGCCA
 CCACACCTGGGTGATTTTTGTATTTTTTAGTAGAGACGGGGTTTCACCATGTTGGTCAGGCTG
 GTCTCAAACCTCCTGACCTAGTGATCCACCCTCCTCGGCCTCCCAAAGTGCTGGGATTACAGG
 CATGAGCCACCACAGCTGGCCCCCTTCTGTTTTATGTTTGGTTTTTGAAGAAGGAATGAAGTG
 GGAACCAAATTAGGTAATTTTTGGGTAATCTGTCTCTAAATATATTAGCTAAAAACAAAGCTCT
 ATGTAAAGTAATAAAGTATAATTGCCATATAAATTTCAAATTCAACTGGCTTTTATGCAAA
 GAAACAGGTTAGGACATCTAGGTTCCAATTCATTCACATTCTTGGTTCCAGATAAAATCAAC
 TGTTTATATCAATTTCTAATGGATTTGCTTTTCTTTTATATGGATTCTTTAAACCTTATT
 CCAGATGTAGTTCCTTCCAATTAAATATTTGAATAAATCTTTTGTACTCAA

FIGURE 328

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA45410

><subunit 1 of 1, 431 aa, 1 stop

><MW: 46810, pI: 6.45, NX(S/T): 6

MFFGGEGSLTYTLVIICFLTLRLSASQNCLKKSLEDVVIDIQSSLKGIKRGNEPVYTSTQED
CINSCCSTKNISGDKACNLMIFDTRKTARQPNCYLFFCPNEEACPLKPAKGLMSYRIITDFP
SLTRNLPSQELPQEDSLLHGQFSQAVTPLAHHHTDYSKPTDISWRDTLSQKFGSSDHLEKLF
KMDEASAQLLAYKEKGHSQSSQFSSDQEIAHLLPENVSALPATVAVASPHHTSATPKPATLL
PTNASVTPSGTSQPQLATTAPPVTTVTSPPTTLISTVFTRAAATLQAMATTAVLTTTFQAP
TDSKGSLETIPFTEISNLTNTGNVYNPTALSMNVESSTMNKTASWEGREASPGSSSQGSV
PENQYGLPFEKWLLIGSLLFGVLFLVIGLVLLGRILSESLRRKRYSRLDYLINGIYVDI

Signal sequence.

amino acids 1-25

Transmembrane domain.

amino acids 384-405

N-glycosylation sites.

amino acids 72-76, 222-226, 251-255, 327-331, 352-356

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 415-419

Tyrosine kinase phosphorylation site.

amino acids 50-57

N-myristoylation sites.

amino acids 4-10, 48-54, 315-321

CTCCACGGTGTCCAGCGCCAGAAATGCGGCTTCTGGTCTGCTATGGGGTTGCCTGCTGCTCTCCAGGTTATGAAGCCCTGGAGGGCCAGAGGAAATCAGCGGGTTTGAAGGGGACACTGTGTCCCTGCAGTGCACCTACAGGGAAGAGCTGAGGGACCACCGGAAGTACTGGTGCAGGAAGGGTGGGATCCTCTTCTCTCGCTGCTCTGGCACCATCTATGCAGAAGAAGAAGGCCAGGAGACAATGAAGGGCAGGGTGTCCATCCGTGACAGCCGCCAGGAGCTCTCGCTCATTGTGACCCTGTGGAACCTCACCCCTGCAAGACGCTGGGGAGTACTGGTGTGGGGTCGAAAAACGGGGCCCCGATGAGTCTTTTACTGATCTCTCTGTTTCGTCTTTCCAGGACCCTGCTGTCTCTCCCTCCCCTTCTCCCACCTTCCAGCCTCTGGCTACAACACGCCTGCAGCCCAAGGCAAAGCTCAGCAAACCCAGCCCOCAGGATTGACTTCTCCTGGGCTCTACCCGGCAGCCACCACAGCCAAGCAGGGGAAGACAGGGGCTGAGAGCCCCCTCCATTGCCAGGGACTTCCCAGTACGGGCACGAAAGGACTTCTCAGTACACAGGAACCTCTCCTCACCCAGCGACCTCTCCTCCTGCAGGGAGCTCCCGCCCCCCCATGCAGCTGGACTCCACCTCAGCAGAGGACACCAGTCCAGCTCTCAGCAGTGGCAGCTCTAAGCCCAGGGTGTCCATCCCAGATGGTCCGCATACTGGCCCCAGTCTTGGTGCTGCTGAGCCTTCTGTGAGCGCAGGCCTGATCGCCTTCTGCAGCCACCTGCTCCTGTGGAGAAAGGAAGCTCAACAGGCCACGGAGACACAGAGGAACGAGAAAGTTCTGGCTCTCACGCTTGACTGCGGAGGAAAAGGAAGGCCCTTCCCAGGCCCCCTGAGGGGGACGTGATCTCGATGCCTCCCCTCCACACATCTGAGGAGGAGCTGGGCTTCTCGAAGTTTGTCTCAGCGTAGGGCAGGAGGCCCTCCTGGCCAGGCCAGCAGTGAAGCAGTATGGCTGGCTGGATCAGCACCGATTCCCGAAAGCTTTCCACCTCAGCCTCAGAGTCCAGCTGCCCGGACTCCAGGGCTCTCCCCACCCTCCCCAGGCTCTCCTCTTGATGTTCCAGCCTGACCTAGAAGCGTTTGTCTAGCCCTGGAGCCCAGAGCGGTGGCCTTGCTCTTCCGGCTGGAGACTGGGACATCCCTGATAGGTTACATCCCTGGGCAGAGTACCAGGCTGCTGACCCTCAGCAGGGCCAGACAAGGCTCAGTGGATCTGGTCTGAGTTTCAATCTGCCAGGAACCTGGGCTCATGCCCAGTGTGCGACCCCTGCCTTCTTCCACTCCAGACCCACCTTGTCTTCCCTCCCTGGCGTCTCTCAGACTTAGTCCCACGGTCTCCTGCATCAGCTGGTGATGAAGAGGAGCATGCTGGGGTGAGACTGGGATTCTGGCTTCTCTTTGAACCACCTGCATCCAGCCCTTCAGGAAGCCTGTGAAAAACGTGATTCTGGCCCCACCAAGACCCACCAAACCATCTCTGGGCTTGGTGCAGGACTCTGAATTCTAACAAATGCCAGTGACTGTGCGACTTGAGTTTGAGGGCCAGTGGGCCTGATGAACGCTCACACCCCTTCAGCTTAGAGTCTGCATTTGGGCTGTGACGTCTCCACCTGCCC CAATAGATCTGCTCTGTCTGCGACACCAGATCCACGTGGGGACTCCCCTGAGGCCTGCTAAGTCCAGGCCTTGGTCAAGTCAAGTGCACATTGCAGGATAAGCCCAGGACCGGCACAGAAGTGGTTGCCTTTNCCATTTGCCCTCCCTGGNCCATGCCTTCTTGCCCTTGGAAAAAATGATGAAGAAACCTTGGCTCCTTCCTTGTCTGGAAAGGGTTACTTGCCTATGGGTTCTGGTGGCTAGAGAGAAAAGTAGAAAACCAGAGTGCACGTAGGTGTCTAACACAGAGGAGAGTAGGAACAGGGCGGATACCTGAAGGTGACTCCGAGTCCAGCCCCCTGGAGAAGGGGTGCGGGGTGGTGGTAAAGTAGCACAACCTACTATTTTTTTTTCTTTTTTCCATTATTATTGTTTTTTAAGACAGAATCTCGTGCTGCTGCCCAGGCTGGAGTGCAGTGGCACGATCTGCAAACCTCCGCCTCCTGGGTTCAAGTGATTCTTCTGCCTCAGCCTCCCAGTAGCTGGGATTACAGGCACGCACCACCACACCTGGCTAATT TTTGTACTTTTAGTAGAGATGGGGTTTACCATGTTGGCCAGGCTGGTCTTGAACCTCCTGACCTCAAATGAGCCTCCTGCTTCAGTCTCCCAAATTGCCGGGATTACAGGCATGAGCCACTGTGTCTGGCCCTATTTCCCTTTAAAAAGTGAAATTAAGAGTTGTTTCAGTATGCAAACTTGGAAGATGGAGGAGAAAAAGAAAAGGAAGAAAAAAATGTCACCCATAGTCTCACCAGAGACTATCAT TATTTTCGTTTTTGTGTACTTCCTTCCACTCTTTTCTTCTTACATAATTTGCCGGTGTTCTTTTACAGAGCAATTATCTTGTATATACAACCTTTGTATCCTGCCTTTTCCACCTTATCGTTCCATCACTTTATTCCAGCACTTCTCTGTGTTTTTACAGACCTTTTTTATAAATAAAATGTTTCATCAGCTGCATAAAAAAAAAAAAAA

FIGURE 330

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA44196

<subunit 1 of 1, 332 aa, 1 stop

<MW: 36143, pI: 5.89, NX(S/T): 1

MRLLVLLWGCLLLPGYEALEGPEEISGFEGDTVSLQCTYREELRDHRKYWCRKGGILFSRCS
GTIYAE EEGQETMKGRVSIRDSRQELSLIVTLWNLTLQDAGEYWCGVEKRGPD E S L L I S L F V
FPGPCCPPSPSPTFQPLATTRLQPKAKAQQTQPPGLTSPGLYPAATTAKQGKTGAEAPPLPG
TSQYGHERTSQYTGTSPHPATSPAGSSRPPMQLDSTSAEDTSPALSSGSSKPRVSIPMVRI
LAPVLVLLSLLSAAGLIAFCSHLLLRKEAQQATETQRNEKFWLSRLTAE E K E A P S Q A P E G D
VISMPPLHTSEEEELGFSKFVSA

Important features:**Signal peptide:**

amino acids 1-17

Transmembrane domain:

amino acids 248-269

N-glycosylation site.

amino acids 96-99

Fibrinogen beta and gamma chains C-terminal domain.

amino acids 104-113

Ig like V-type domain:

amino acids 13-128